GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-12
Perfect score: 20
Sequence: 1 cgaccccacagacagccccc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.8	74.0	41	12	BG616493	BG616493 602614656
	2	14.2	71.0	85	17	AZ428622	AZ428622 1M0212N10
c	3	13.8	69.0	40	9	AA877262	AA877262 nq94d06.s
	4	13.8	69.0	49	9	AI815199	AI815199 wk72c04.x
c	5	13.8	69.0	50	9	AU104946	AU104946 AU104946
	6	13.8	69.0	76	9	AI601162	AI601162 ar89e02.x
	7	13.8	69.0	79	14	H42857	H42857 yo10e01.s1
c	8	13.8	69.0	93	17	AZ949869	AZ949869 2M0213I07
	9	13.8	69.0	100	12	BG213902	BG213902 RST33525
	10	13.6	68.0	61	9	AI015169	AI015169 ot74g05.s
	11	13.6	68.0	64	9	AI061438	AI061438 an29a06.x
	12	13.6	68.0	66	17	AQ254888	AQ254888 EP(3)0464
	13	13.6	68.0	70	9	AI856414	AI856414 sb41h10.x
	14	13.6	68.0	73	17	AQ254591	AQ254591 EP(3)0593
c	15	13.6	68.0	82	14	T25069	T25069 EST644 Huma
c	16	13.6	68.0	92	9	AA886273	AA886273 oj38f06.s
c	17	13.4	67.0	88	9	AA270168	AA270168 va58h01.r
c	18	13.4	67.0	97	14	BQ614622	BQ614622 ik99e01.y
c	19	13.2	66.0	45	17	AZ954608	AZ954608 2M0220B24
	20	13.2	66.0	61	9	AI612903	AI612903 ty33d04.x
	21	13.2	66.0	61	14	H39535	H39535 yl68a08.r1
	22	13.2	66.0	66	17	AF219041	AF219041 AF219041
c	23	13.2	66.0	72	14	BQ271752	BQ271752 ik17h10.x
c	24	13.2	66.0	73	13	BM128826	BM128826 if17b08.x
c	25	13.2	66.0	81	9	AI828860	AI828860 ts81g12.x
	26	13.2	66.0	84	10	BE566379	BE566379 601339975
	27	13.2	66.0	84	14	W21701	W21701 zEST00787 M
c	28	13.2	66.0	86	13	BM312177	BM312177 ig41b10.x
c	29	13.2	66.0	90	13	BM353195	BM353195 ig76h05.x
c	30	13.2	66.0	97	10	BE463686	BE463686 hx79b04.x
c	31	13.2	66.0	98	13	BI324964	BI324964 ic20d11.x
c	32	13.2	66.0	99	10	AW583476	AW583476 ia05d12.x
c	33	13.2	66.0	99	13	BM272195	BM272195 ig40d11.x
	34	13.2	66.0	99	14	BQ272071	BQ272071 ik17h10.y
	35	13.2	66.0	99	17	BH230402	BH230402 1006157G0
c	36	12.8	64.0	34	17	AZ387869	AZ387869 1M0147N23
c	37	12.8	64.0	52	17	AZ511071	AZ511071 1M0355M21
	38	12.8	64.0	73	9	AI941442	AI941442 sb89h10.y
c	39	12.8	64.0	74	9	AI274464	AI274464 qv61b05.x
c	40	12.8	64.0	74	17	AZ433977	AZ433977 1M0220G09
	41	12.8	64.0	79	9	AI571011	AI571011 to25d09.x
	42	12.8	64.0	83	14	BQ764538	BQ764538 EBca01_SQ
	43	12.8	64.0	87	14	F32349	F32349 HSPD25000 H
c	44	12.8	64.0	94	13	BM310585	BM310585 ig50a02.x
c	45	12.8	64.0	100	17	AZ819462	AZ819462 2M0091A10

ALIGNMENTS

RESULT 1
BG616493
LOCUS BG616493 41 bp mRNA linear EST 18-APR-2001
DEFINITION 602614656F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4733796 5', mRNA sequence.
ACCESSION BG616493
VERSION BG616493.1 GI:13667864
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 41)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1595 row: o column: 13.
FEATURES Location/Qualifiers
source 1. .41
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="IMAGE:4733796"
/clone_lib="NIH_MGC_76"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.85 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

BASE COUNT 10 a 23 c 5 g 3 t
ORIGIN

Query Match 74.0%; Score 14.8; DB 12; Length 41;
Best Local Similarity 88.9%; Pred. No. 1.7e+04;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       3 ACCCCACAGACAGCCCCC 20
           |||||||  |||||||||
Db      11 ACCCCACCCACAGCCCCC 28
```

RESULT 2
AZ428622
LOCUS AZ428622 85 bp DNA linear GSS 03-OCT-2000
DEFINITION 1M0212N10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0212N10 F, DNA sequence.
ACCESSION AZ428622
VERSION AZ428622.1 GI:10552635
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 85)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0212 row: N column: 10
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 85.
FEATURES Location/Qualifiers
source 1. .85
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0212N10"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 18 a 37 c 15 g 15 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 17; Length 85;
Best Local Similarity 84.2%; Pred. No. 3.3e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 GACCCCACAGACAGCCCCC 20
           || | ||||||| |||||
Db      15 GAACTCACAGACACCCCCC 33
```

RESULT 3
AA877262/c
LOCUS AA877262 40 bp mRNA linear EST 31-MAR-1998
DEFINITION nq94d06.s1 NCI_CGAP_Co10 Homo sapiens cDNA clone IMAGE:1159979 3' similar to gb:M16038 TYROSINE-PROTEIN KINASE LYN (HUMAN);, mRNA sequence.
ACCESSION AA877262
VERSION AA877262.1 GI:2986339
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 40)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 1388 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .40
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1159979"
/clone_lib="NCI_CGAP_Co10"
/tissue_type="colon tumor RER+"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from RER+ colon tumor, and was then primed with a Not I -oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3

vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo (N-Soares4). "
BASE COUNT 4 a 11 c 18 g 7 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 40;
Best Local Similarity 88.2%; Pred. No. 4.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACCCCACAGACAGCC 17
||| |||| ||||||||
Db 38 CGAGCCCAGAGACAGCC 22

RESULT 4
AI815199
LOCUS AI815199 49 bp mRNA linear EST 07-MAR-2000
DEFINITION wk72c04.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2420934 3' similar to SW:PRP4_HUMAN P10163 SALIVARY PROLINE-RICH PROTEIN PO PRECURSOR ;, mRNA sequence.
ACCESSION AI815199
VERSION AI815199.1 GI:5426404
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 49)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 300 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .49
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2420934"
/clone_lib="NCI_CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"
BASE COUNT 17 a 22 c 6 g 4 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 49;
Best Local Similarity 88.2%; Pred. No. 4.5e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CCCCACAGACAGCCCCC 20
||||| ||| |||||||
Db 26 CCCCAAAGAAAGCCCCC 42

RESULT 5
AU104946/c
LOCUS AU104946 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU104946 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HRC05670, mRNA sequence.
ACCESSION AU104946
VERSION AU104946.1 GI:13554467
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HRC05670"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and dimethylfumarate treated U937 cells"
BASE COUNT 4 a 10 c 29 g 7 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 50;
Best Local Similarity 88.2%; Pred. No. 4.5e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CCCCACAGACAGCCCCC 20
|||| |||||| |||||
Db 38 CCCCGCAGACACCCCCC 22

RESULT 6
AI601162
LOCUS AI601162 76 bp mRNA linear EST 21-APR-1999
DEFINITION ar89e02.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone IMAGE:2152442 3' similar to TR:Q62901 Q62901 ATROPHIN-1 RELATED PROTEIN. ;, mRNA sequence.
ACCESSION AI601162
VERSION AI601162.1 GI:4610368
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 76)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand

Seq primer: -40UP from Gibco
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .76
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2152442"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATTCACTAGTAAT 3' and 5' ATTACTAGTG 3'], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead."
BASE COUNT 15 a 38 c 11 g 12 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 76;
Best Local Similarity 88.2%; Pred. No. 4.8e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CCCCACAGACAGCCCCC 20
||||||||| || ||||
Db 27 CCCCACAGATAGACCCC 43

RESULT 7
H42857
LOCUS H42857 79 bp mRNA linear EST 31-JUL-1995
DEFINITION yo10e01.s1 Soares adult brain N2b5HB55Y Homo sapiens cDNA clone IMAGE:177528 3' similar to gb:M64241 QM PROTEIN (HUMAN);, mRNA sequence.
ACCESSION H42857
VERSION H42857.1 GI:918909
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 79)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1472
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 1472 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .79
/organism="Homo sapiens"
/db_xref="GDB:3839724"
/db_xref="taxon:9606"
/clone="IMAGE:177528"
/clone_lib="Soares adult brain N2b5HB55Y"
/sex="Male"
/dev_stage="55-year old"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: brain; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 53. Library constructed by Bento Soares and M.Fatima Bonaldo. The adult brain RNA was provided by Dr. Donald H. Gilden. Tissue was acquired 17-18 hours after death which occurred in consequence of a ruptured aortic aneurysm. RNA was prepared from a pool of tissues representing the following areas of the brain: frontal, parietal, temporal and occipital cortex from the left and right hemispheres, subcortical white matter, basal ganglia, thalamus, cerebellum, midbrain, pons and medulla."
BASE COUNT 16 a 26 c 18 g 11 t 8 others
ORIGIN

Query Match 69.0%; Score 13.8; DB 14; Length 79;
Best Local Similarity 88.2%; Pred. No. 4.8e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GACCCCACAGACAGCCC 18
||||||||||| || |||
Db 28 GACCCCACAGCCATCCC 44

RESULT 8
AZ949869/c
LOCUS AZ949869 93 bp DNA linear GSS 27-APR-2001
DEFINITION 2M0213I07R Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0213I07 R, DNA sequence.
ACCESSION AZ949869
VERSION AZ949869.1 GI:13821096
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 93)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0213 row: I column: 07
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 93.
FEATURES Location/Qualifiers
source 1. .93
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"

/clone="UUGC2M0213I07"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 14 a 14 c 34 g 31 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 93;
Best Local Similarity 88.2%; Pred. No. 4.9e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ACCCCACAGACAGCCCC 19
|||||| | ||||||||
Db 71 ACCCCATATACAGCCCC 55

RESULT 9
BG213902
LOCUS BG213902 100 bp mRNA linear EST 21-APR-2001
DEFINITION RST33525 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG213902
VERSION BG213902.1 GI:13735589
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 100)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com.
FEATURES Location/Qualifiers
source 1. .100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."
BASE COUNT 32 a 31 c 20 g 17 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 12; Length 100;
Best Local Similarity 88.2%; Pred. No. 4.9e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ACCCCACAGACAGCCCC 19
|||||| ||| ||||||
Db 46 ACCCCAAAGAGAGCCCC 62

RESULT 10
AI015169
LOCUS AI015169 61 bp mRNA linear EST 16-JUN-1998
DEFINITION ot74g05.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:1622552 3' similar to TR:P94915 P94915 GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE HOMOLOG. ;, mRNA sequence.
ACCESSION AI015169
VERSION AI015169.1 GI:3229505
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 61)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .61
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1622552"
/clone_lib="Soares_total_fetus_Nb2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCTTAATTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 12 a 25 c 21 g 3 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 61;
Best Local Similarity 80.0%; Pred. No. 5.6e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGACCCCACAGACAGCCCCC 20
|| | |||||| | ||||||
Db 3 CGGCACCACAGCCGGCCCCC 22

RESULT 11
AI061438
LOCUS AI061438 64 bp mRNA linear EST 22-JUL-1998
DEFINITION an29a06.x1 Gessler Wilms tumor Homo sapiens cDNA clone IMAGE:1700050 3' similar to SW:CA1B_HUMAN P12107 COLLAGEN ALPHA

```
          1(XI) CHAIN PRECURSOR. [1] ;contains MSR1.t1 MSR1 repetitive
          element ;, mRNA sequence.
ACCESSION AI061438
VERSION   AI061438.1  GI:3336806
KEYWORDS  EST.
SOURCE    human.
  ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1  (bases 1 to 64)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
          Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
          ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
          White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT   Contact: Wilson RK
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: est@watson.wustl.edu
          This clone is available royalty-free through LLNL ; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          Trace considered overall poor quality
          Possible reversed clone: similarity on wrong strand
          Seq primer: -40m13 fwd. ET from Amersham
          High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1700050"
                     /clone_lib="Gessler Wilms tumor"
                     /sex="pooled (6)"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; RNA
                     was prepared from a pool of 6 anonymous Wilms' tumor RNAs.
                     RNA was prepared by acid-phenol, followed by one round of
                     oligo dT selection.  cDNA library preparation was with
                     the BRL/Life Tech. Superscript Plasmid system.  An
                     oligo-dT NotI primer for first strand synthesis generated
                     gcggccgccc(t)n at the 3' end of the clones.  A 5' SalI
                     adaptor was used  with sequence 5'-gtcgacccacgcgtccg-3'.
                     Resulting cDNAs were size selected (average size 2 kb),
                     NotI digested, and ligated into NotI/SalI-cut pSPORT1.
                     Library was constructed by Dr. Manfred Gessler."
BASE COUNT       15 a     28 c     21 g      0 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 64;
  Best Local Similarity   80.0%;  Pred. No. 5.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           || |||| || | |||||||
Db       3 CGCCCCCCCAAAAAGCCCCC 22


RESULT 12
AQ254888
LOCUS       AQ254888                 66 bp    DNA     linear   GSS 23-AUG-2000
DEFINITION  EP(3)0464-3prime Drosophila melanogaster EP line Drosophila
          melanogaster genomic Sequence recovered from 3' end of P element,
          DNA sequence.
ACCESSION   AQ254888
VERSION     AQ254888.1  GI:3738446
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1  (bases 1 to 66)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
          Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
COMMENT   Contact: Gerald Rubin
          Berkeley Drosophila Genome Project
          University of California, Berkeley
          LSA Building, Berkeley, CA 94720-3200, USA
          Fax: 5106439947
          Email: gerry@fruitfly.berkeley.edu
          Sequence recovery method was inverse PCR.

          Sequence orientation is forward strand relative to 5' end of P
          element

          The P element insertion position is base 001 in the 66 bases. This
          insertion position refers to the first base of the 8 base target
          recognition sequence.
          Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster EP line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single EP
                     transposable element insertion. (The generation of these
                     insertion strains is described in Rorth P, Szabo K, Bailey
                     A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
                     V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function
                     genetics in Drosophila. Development 6:1049-1057.) The
                     resultant fragment for each strain was directly sequenced
                     to determine the genomic sequence at the site of
                     insertion. Details of the protocols used can be found at
                     http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT       15 a     23 c     17 g     11 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 66;
  Best Local Similarity   80.0%;  Pred. No. 5.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           |||||||||| |  | |||
Db      38 CGACCCCACAGCCCCCGCCC 57


RESULT 13
AI856414
LOCUS       AI856414                 70 bp    mRNA    linear   EST 30-NOV-2001
DEFINITION  sb41h10.x1 Gm-c1014 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
          Gm-c1014-284 3' similar to TR:Q40385 Q40385 PISTIL EXTENSIN-LIKE
          PROTEIN. ;, mRNA sequence.
ACCESSION   AI856414
VERSION     AI856414.1  GI:5509856
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
          Glycine.
REFERENCE 1  (bases 1 to 70)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
          ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
          Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
          ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
          ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
          ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Other_ESTs: sb41h10.y1
            Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand This clone is
            available through: ResGen, Invitrogen Corp. 2130 South Memorial
            Parkway Huntsville, AL 35801 For further information call: (800
            )-533-4363 or contact via email: ccu@resgen.com
            Seq primer: primer name ambiguous
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1014-284"
                     /clone_lib="Gm-c1014"
                     /tissue_type="Leaves, 2-3 week old seedlings, greenhouse
                     grown"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3Pac (pT7T3, Pharmacia); Site_1: EcoRI;
                     Site_2: HindIII; This cDNA library was constructed from
                     mRNA isolated from leaves of 2-3 week old greenhouse grown
                     plants. Complementary DNA was synthesized from mRNA using
                     a 3' anchored poly (dT) primer. EcoRI adapters were
                     ligated to the blunt-ended cDNA fragments followed by
                     digestion with EcoRI and HindIII. The cDNA fragments were
                     directionally cloned into the EcoRI-HindIII restriction
                     site of the pT7T3-Pac vector. The ligated cDNA fragments
                     were transformed into DH10B host cells (Gibco BRL). This
                     library was constructed by Dr. Randy Shoemaker and Dr.
                     John Erpelding."
BASE COUNT        9 a     33 c     27 g      1 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 70;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | |||||||| ||| ||| |
Db     43 CCACCCCACACACACCCCGC 62


RESULT 14
AQ254591
LOCUS       AQ254591                  73 bp    DNA     linear   GSS 23-AUG-2000
DEFINITION  EP(3)0593-3prime Drosophila melanogaster EP line Drosophila
            melanogaster genomic Sequence recovered from 3' end of P element,
            DNA sequence.
ACCESSION   AQ254591
VERSION     AQ254591.1  GI:3738454
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
            Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 001 in the 73 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster EP line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single EP
                     transposable element insertion. (The generation of these
                     insertion strains is described in Rorth P, Szabo K, Bailey
                     A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
                     V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function
                     genetics in Drosophila. Development 6:1049-1057.) The
                     resultant fragment for each strain was directly sequenced
                     to determine the genomic sequence at the site of
                     insertion. Details of the protocols used can be found at
                     http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT       18 a     25 c     18 g     12 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 73;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          |||||||||||  |  | |||
Db     45 CGACCCCACAGCCCCCGCCC 64


RESULT 15
T25069/c
LOCUS       T25069                    82 bp    mRNA    linear   EST 24-AUG-1995
DEFINITION  EST644 Human colorectal cancer Homo sapiens cDNA clone 30B12, mRNA
            sequence.
ACCESSION   T25069
VERSION     T25069.1  GI:534694
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Frigerio,J.-M., Berthezene,P., Garrido,P., Ortiz,E., Barthellemy,S,
            Vasseur,S., Sastre,B., Seleznieff,I., Dagorn,J.-C. and Iovanna
            ,J.-L.
  TITLE     Analysis of 2166 clones from a human colorectal cancer cDNA library
            by partial sequencing
  JOURNAL   Hum. Mol. Genet. 4, 37-43 (1995)
  MEDLINE   95227175
COMMENT     Contact: Iovanna JL
            U.315 INSERM
            46 Bd de la Gaye, F-13009 Marseille, France.
            Tel: (33) 91 82 03 15
            Fax: (33) 91 26 62 19
            Email: dagorn@arthur.citi2.fr
            This sequence is one of a series obtained by systematic  sequencing
            of a colorectal cancer cDNA library.
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="30B12"
                     /clone_lib="Human colorectal cancer"
```

/lab_host="E. coli NM522"
/note="Vector: pT7T3D; Site_1: Eco RI; Site_2: Not I; mRNA was purified from a colorectal tumour of an adult male. cDNA was constructed and cloned into the pT7T3D phagemid following the manufacturer intructions (Pharmacia)."

BASE COUNT 9 a 23 c 35 g 15 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 14; Length 82;
Best Local Similarity 80.0%; Pred. No. 5.8e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CGACCCCACAGACAGCCCCC 20
           || | |||| |||||| |||
Db      64 CGGCTCCACGGACAGCACCC 45
```

Search completed: November 26, 2002, 04:08:13
Job time : 766.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-12
Perfect score: 20
Sequence: 1 cgaccccacagacagccccc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		13.2	66.0	90	10	US-09-864-761-19125	Sequence 19125, A
c 2		12.8	64.0	80	10	US-09-764-877-400	Sequence 400, App
c 3		12.8	64.0	82	9	US-10-079-623-80	Sequence 80, Appl
c 4		12.6	63.0	25	10	US-09-866-108-5565	Sequence 5565, Ap
c 5		12.6	63.0	25	10	US-09-866-108-5566	Sequence 5566, Ap
c 6		12.6	63.0	25	10	US-09-866-108-5567	Sequence 5567, Ap
c 7		12.6	63.0	25	10	US-09-866-108-5568	Sequence 5568, Ap
c 8		12.6	63.0	25	10	US-09-866-108-5569	Sequence 5569, Ap
c 9		12.6	63.0	25	10	US-09-866-108-5570	Sequence 5570, Ap
c 10		12.6	63.0	25	10	US-09-866-108-5571	Sequence 5571, Ap
c 11		12.6	63.0	31	10	US-09-374-671-77	Sequence 77, Appl
c 12		12.6	63.0	38	10	US-09-884-877-4	Sequence 4, Appli
13		12.6	63.0	40	10	US-09-931-184-6	Sequence 6, Appli
14		12.6	63.0	72	10	US-09-922-261-118	Sequence 118, App
15		12.6	63.0	90	10	US-09-922-261-116	Sequence 116, App
c 16		12.2	61.0	17	10	US-09-866-108-2636	Sequence 2636, Ap
17		12.2	61.0	20	10	US-09-350-259-79	Sequence 79, Appl
c 18		12.2	61.0	25	10	US-09-866-108-5572	Sequence 5572, Ap
c 19		12.2	61.0	25	10	US-09-866-108-5573	Sequence 5573, Ap
c 20		12.2	61.0	26	10	US-09-350-259-67	Sequence 67, Appl
c 21		12.2	61.0	31	10	US-09-801-274-410	Sequence 410, App
c 22		12.2	61.0	89	10	US-09-864-761-20205	Sequence 20205, A
c 23		12.2	61.0	100	9	US-10-047-542-40	Sequence 40, Appl
c 24		12	60.0	18	10	US-09-768-436-34	Sequence 34, Appl
c 25		12	60.0	93	10	US-09-878-574-9664	Sequence 9664, Ap
26		11.8	59.0	23	10	US-09-216-393-220	Sequence 220, App
27		11.8	59.0	23	10	US-09-832-723-2	Sequence 2, Appli
28		11.8	59.0	25	10	US-09-866-073-8	Sequence 8, Appli
c 29		11.8	59.0	31	10	US-09-801-274-237	Sequence 237, App
c 30		11.8	59.0	47	9	US-09-853-526-282	Sequence 282, App
c 31		11.8	59.0	47	10	US-09-901-484A-282	Sequence 282, App
32		11.8	59.0	100	10	US-09-983-965-4933	Sequence 4933, Ap
c 33		11.6	58.0	20	10	US-09-735-995-92	Sequence 92, Appl
c 34		11.6	58.0	25	10	US-09-866-108-5564	Sequence 5564, Ap
35		11.6	58.0	31	10	US-09-801-274-1698	Sequence 1698, Ap
c 36		11.6	58.0	80	10	US-09-864-761-22206	Sequence 22206, A
c 37		11.6	58.0	85	10	US-09-864-761-23386	Sequence 23386, A
c 38		11.6	58.0	97	10	US-09-867-701-4678	Sequence 4678, Ap
39		11.4	57.0	20	10	US-09-749-728B-60	Sequence 60, Appl
c 40		11.4	57.0	31	10	US-09-801-274-238	Sequence 238, App
c 41		11.4	57.0	35	9	US-09-453-289-1	Sequence 1, Appli
c 42		11.4	57.0	35	10	US-09-214-606-1	Sequence 1, Appli
c 43		11.4	57.0	35	12	US-10-036-507-16	Sequence 16, Appl
44		11.4	57.0	41	10	US-09-828-366-19	Sequence 19, Appl
c 45		11.4	57.0	89	10	US-09-861-893-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-09-864-761-19125
; Sequence 19125, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19125
;   LENGTH: 90
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007425.16
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.97
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
;   OTHER INFORMATION: EST_HUMAN HIT: AW172350.1, EVALUE 1.20e-02
;   OTHER INFORMATION: NT HIT: X03638.1, EVALUE 1.90e-01
US-09-864-761-19125

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 90;
  Best Local Similarity   83.3%;  Pred. No. 1.3e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACCCCACAGACAGCCCC 19
          ||  || ||||||||||||
Db      9 GAATCCCCAGACAGCCCC 26


RESULT 2
US-09-764-877-400/c
; Sequence 400, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005
;  CURRENT APPLICATION NUMBER: US/09/764,877
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 400
;   LENGTH: 80
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-877-400

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 80;
  Best Local Similarity   87.5%;  Pred. No. 2e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 CCCACAGACAGCCCCC 20
          |||||||||| ||| |
Db     51 CCCACAGACAACCCAC 36


RESULT 3
US-10-079-623-80/c
; Sequence 80, Application US/10079623
; Patent No. US20020169302A1
; GENERAL INFORMATION:
;  APPLICANT: Havukkala, Ilkka J.
;  APPLICANT:  Glenn, Matthew
;  APPLICANT:  Grigor, Murray R.
;  APPLICANT:  Molenaar, Adrian J.
;  TITLE OF INVENTION: Compositions isolated from bovine
;  TITLE OF INVENTION:  mammary gland and methods for their use.
;  FILE REFERENCE: 11000.1044c3
;  CURRENT APPLICATION NUMBER: US/10/079,623
;  CURRENT FILING DATE:  2002-02-19
;  NUMBER OF SEQ ID NOS: 370
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
;   LENGTH: 82
;   TYPE: DNA
;   ORGANISM: Bovine
US-10-079-623-80

  Query Match             64.0%;  Score 12.8;  DB 9;  Length 82;
  Best Local Similarity   87.5%;  Pred. No. 2e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGC 16
          |||||||| |||| ||
Db     18 CGACCCCATAGACGGC 3


RESULT 4
US-09-866-108-5565/c
; Sequence 5565, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
```

```
; SEQ ID NO 5565
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-5565

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity  78.9%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GACCCCACAGACAGCCCCC 20
           |||||  || ||||||| |
Db      25 GACCCGCCAAACAGCCCGC 7


RESULT 5
US-09-866-108-5566/c
; Sequence 5566, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 5566
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-5566

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity  78.9%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GACCCCACAGACAGCCCCC 20
           |||||  || ||||||| |
Db      24 GACCCGCCAAACAGCCCGC 6


RESULT 6
US-09-866-108-5567/c
; Sequence 5567, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 5567
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-5567

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity  78.9%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GACCCCACAGACAGCCCCC 20
           |||||  || ||||||| |
Db      23 GACCCGCCAAACAGCCCGC 5


RESULT 7
US-09-866-108-5568/c
; Sequence 5568, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 5568
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-5568

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GACCCCACAGACAGCCCCC 20
           |||||  || |||||||  |
Db      22 GACCCGCCAAACAGCCCGC 4


RESULT 8
US-09-866-108-5569/c
; Sequence 5569, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 5569
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-5569

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GACCCCACAGACAGCCCCC 20
           |||||  || |||||||  |
Db      21 GACCCGCCAAACAGCCCGC 3


RESULT 9
US-09-866-108-5570/c
; Sequence 5570, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 5570
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-5570

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GACCCCACAGACAGCCCCC 20
          |||||  || ||||||| |
Db     20 GACCCGCCAAACAGCCCGC 2


RESULT 10
US-09-866-108-5571/c
; Sequence 5571, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 5571
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-5571

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GACCCCACAGACAGCCCCC 20
          |||||  || ||||||| |
Db     19 GACCCGCCAAACAGCCCGC 1


RESULT 11
US-09-374-671-77/c
; Sequence 77, Application US/09374671
; Patent No. US20020012963A1
;  GENERAL INFORMATION:
;    APPLICANT:  Morgenstern, Jay P.
;    APPLICANT:  Kanieczny, Andrzey
;    APPLICANT:  Bizindaukas, Christine B.
;    APPLICANT:  Brauer, Andrew W.
;    TITLE OF INVENTION:  Allergenic Proteins and Peptides from Dog
;    TITLE OF INVENTION:  Dander and Uses Therefor
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII-text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/374,671
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/491,861
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  IMI-026CP(IPC-048CP)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 742-4214
;  INFORMATION FOR SEQ ID NO:  77:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
```

US-09-374-671-77

Query Match 63.0%; Score 12.6; DB 10; Length 31;
Best Local Similarity 78.9%; Pred. No. 2.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CGACCCCACAGACAGCCCC 19
           | ||| | |||||||| ||
Db      31 CCACCTCCCAGACAGCTCC 13
```

RESULT 12
US-09-884-877-4/c
; Sequence 4, Application US/09884877
; Patent No. US20020151059A1
; GENERAL INFORMATION:
; APPLICANT: TE RIELE, HENRICUS
; APPLICANT: DE WIND, NIELS
; APPLICANT: DEKKER-VLAAR, HELENA
; TITLE OF INVENTION: HOMOLOGOUS RECOMBINATION IN MISMATCH REPAIR INACTIVATED
; TITLE OF INVENTION: EUKARYOTIC CELLS
; FILE REFERENCE: 065691/0230
; CURRENT APPLICATION NUMBER: US/09/884,877
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 09/147,712
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/EP95/02980
; PRIOR FILING DATE: 1995-07-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-884-877-4

Query Match 63.0%; Score 12.6; DB 10; Length 38;
Best Local Similarity 78.9%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       2 GACCCCACAGACAGCCCCC 20
           | |||||| | ||||| ||
Db      26 GGCCCCACTGTCAGCCTCC 8
```

RESULT 13
US-09-931-184-6
; Sequence 6, Application US/09931184
; Patent No. US20020119095A1
; GENERAL INFORMATION:
; APPLICANT: Gabathuler, Reinhard
; APPLICANT: Kolaitis, Gerrassimos
; APPLICANT: Brooks, Robert
; APPLICANT: Chen, Qingqi
; APPLICANT: Karkan, Delara
; APPLICANT: Arthur, Gavin
; APPLICANT: St. Pierre, Jean Paul
; APPLICANT: Jeffries, Wilfred
; APPLICANT: Vitalis, Timothy
; TITLE OF INVENTION: Chemotherapeutic Agents Conjugated to P97 and Their Methods of Us
; TITLE OF INVENTION: Treating Neurological Tumours
; FILE REFERENCE: 7685-47
; CURRENT APPLICATION NUMBER: US/09/931,184
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/226,254
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide WJ47
US-09-931-184-6

Query Match 63.0%; Score 12.6; DB 10; Length 40;
Best Local Similarity 78.9%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CGACCCCACAGACAGCCCC 19
           |||  || ||| |||||||
Db      13 CGAGGCCCCAGCCAGCCCC 31
```

RESULT 14
US-09-922-261-118
; Sequence 118, Application US/09922261
; Patent No. US20020111471A1
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/922,261
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/09/461,697
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 72
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-261-118

Query Match 63.0%; Score 12.6; DB 10; Length 72;
Best Local Similarity 78.9%; Pred. No. 2.4e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CGACCCCACAGACAGCCCC 19
           | || |||| ||||||| |
Db       8 CCACTCCACTGACAGCCTC 26
```

RESULT 15
US-09-922-261-116
; Sequence 116, Application US/09922261
; Patent No. US20020111471A1
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/922,261
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/09/461,697
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 116
; LENGTH: 90

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-922-261-116

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 90;
  Best Local Similarity   78.9%;  Pred. No. 2.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCC 19
           | || |||| ||||||| |
Db      26 CCACTCCACTGACAGCCTC 44


Search completed: November 23, 2002, 06:42:09
Job time : 17.25 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-12
Perfect score: 20
Sequence: 1 cgaccccacagacagccccc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.2	71.0	60	2	US-08-353-476-56	Sequence 56, Appl
c	2	13.8	69.0	25	1	US-08-447-179-8	Sequence 8, Appli
	3	13.8	69.0	30	4	US-09-270-542-11	Sequence 11, Appl
c	4	13.6	68.0	26	1	US-08-145-704-20	Sequence 20, Appl
c	5	13.6	68.0	26	1	US-08-145-704-25	Sequence 25, Appl
c	6	13.6	68.0	26	1	US-08-145-704-27	Sequence 27, Appl
c	7	13.6	68.0	26	3	US-08-987-574-20	Sequence 20, Appl
c	8	13.6	68.0	26	3	US-08-987-574-25	Sequence 25, Appl
c	9	13.6	68.0	26	3	US-08-987-574-27	Sequence 27, Appl
c	10	13.6	68.0	26	4	US-08-535-168-20	Sequence 20, Appl
c	11	13.6	68.0	26	4	US-08-535-168-25	Sequence 25, Appl
c	12	13.6	68.0	26	4	US-08-535-168-27	Sequence 27, Appl
c	13	13.6	68.0	26	4	US-09-017-974-20	Sequence 20, Appl
c	14	13.6	68.0	26	4	US-09-017-974-25	Sequence 25, Appl
c	15	13.6	68.0	26	4	US-09-017-974-27	Sequence 27, Appl
c	16	13.6	68.0	26	4	US-08-682-255A-20	Sequence 20, Appl
c	17	13.6	68.0	26	4	US-08-682-255A-25	Sequence 25, Appl
c	18	13.6	68.0	26	4	US-08-682-255A-27	Sequence 27, Appl
c	19	13.6	68.0	26	4	US-09-429-130-20	Sequence 20, Appl
c	20	13.6	68.0	26	4	US-09-429-130-25	Sequence 25, Appl
c	21	13.6	68.0	26	4	US-09-429-130-27	Sequence 27, Appl
c	22	13.6	68.0	26	5	PCT-US96-11786-20	Sequence 20, Appl
c	23	13.6	68.0	26	5	PCT-US96-11786-25	Sequence 25, Appl
c	24	13.6	68.0	26	5	PCT-US96-11786-27	Sequence 27, Appl
c	25	13.6	68.0	45	1	US-08-145-704-21	Sequence 21, Appl
c	26	13.6	68.0	45	1	US-08-145-704-22	Sequence 22, Appl
c	27	13.6	68.0	45	3	US-08-987-574-21	Sequence 21, Appl
c	28	13.6	68.0	45	3	US-08-987-574-22	Sequence 22, Appl
c	29	13.6	68.0	45	4	US-08-535-168-21	Sequence 21, Appl
c	30	13.6	68.0	45	4	US-08-535-168-22	Sequence 22, Appl
c	31	13.6	68.0	45	4	US-09-017-974-21	Sequence 21, Appl
c	32	13.6	68.0	45	4	US-09-017-974-22	Sequence 22, Appl
c	33	13.6	68.0	45	4	US-08-682-255A-21	Sequence 21, Appl
c	34	13.6	68.0	45	4	US-08-682-255A-22	Sequence 22, Appl
c	35	13.6	68.0	45	4	US-09-429-130-21	Sequence 21, Appl
c	36	13.6	68.0	45	4	US-09-429-130-22	Sequence 22, Appl
c	37	13.6	68.0	45	5	PCT-US96-11786-21	Sequence 21, Appl
c	38	13.6	68.0	45	5	PCT-US96-11786-22	Sequence 22, Appl
c	39	13.4	67.0	32	2	US-08-867-941-60	Sequence 60, Appl
c	40	13.4	67.0	32	4	US-09-074-658-60	Sequence 60, Appl
c	41	13.2	66.0	27	2	US-08-356-060A-45	Sequence 45, Appl
c	42	13.2	66.0	27	4	US-08-460-900C-45	Sequence 45, Appl
c	43	13.2	66.0	27	4	US-08-674-509B-45	Sequence 45, Appl
c	44	13.2	66.0	27	4	US-08-954-698-45	Sequence 45, Appl
c	45	13.2	66.0	27	4	US-08-957-874-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-08-353-476-56
; Sequence 56, Application US/08353476
; Patent No. 5871902
; GENERAL INFORMATION:
; APPLICANT: Weininger, Susan
; APPLICANT: Weininger, Arthur M
; TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A
; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st St., Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,476
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bencen, Gerard H
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: GP-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-353-476-56

Query Match 71.0%; Score 14.2; DB 2; Length 60;
Best Local Similarity 84.2%; Pred. No. 4.6e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 GACCCCACAGACAGCCCCC 20
           ||| |||||||||  | |||
```

```
Db      15 GACACCACAGACACCGCCC 33


RESULT 2
US-08-447-179-8/c
; Sequence 8, Application US/08447179
; Patent No. 5744303
;  GENERAL INFORMATION:
;    APPLICANT:  Iggo, Richard
;    APPLICANT:  Friend, Stephen H.
;    APPLICANT:  Frebourg, Thierry
;    APPLICANT:  Ishioka, Chikashi
;    TITLE OF INVENTION:  FUNCTIONAL ASSAY FOR TUMOR
;    TITLE OF INVENTION:  SUPPRESSOR GENES
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,179
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/046,033
;      FILING DATE:  12 APRIL 1993
;      APPLICATION NUMBER:  07/956,696
;      FILING DATE:  10 OCTOBER 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/159002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-447-179-8

  Query Match             69.0%;  Score 13.8;  DB 1;  Length 25;
  Best Local Similarity   88.2%;  Pred. No. 6.7e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCCCACAGACAGCCCCC 20
           |||||| ||||||| ||
Db      20 CCCCACTGACAGCCTCC 4


RESULT 3
US-09-270-542-11
; Sequence 11, Application US/09270542
; Patent No. 6322976
; GENERAL INFORMATION:
;  APPLICANT: Aitman, Timothy
;  APPLICANT: Scott, James
;  APPLICANT: Stanton, Lawrence
;  TITLE OF INVENTION: Compositions and Methods of Disease Diagnosis and
;  TITLE OF INVENTION: Therapy
;  FILE REFERENCE: 4198/78179
;  CURRENT APPLICATION NUMBER: US/09/270,542
;  CURRENT FILING DATE: 1999-03-17
;  EARLIER APPLICATION NUMBER: 09/221,222
;  EARLIER FILING DATE: 1999-12-23
;  NUMBER OF SEQ ID NOS: 207
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-09-270-542-11

  Query Match             69.0%;  Score 13.8;  DB 4;  Length 30;
  Best Local Similarity   88.2%;  Pred. No. 6.7e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GACCCCACAGACAGCCC 18
           ||||||| ||||||| |
Db       9 GACCCCAAAGACAGCAC 25


RESULT 4
US-08-145-704-20/c
; Sequence 20, Application US/08145704
; Patent No. 5567604
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Fennewald, Susan
;    APPLICANT:  Zendegui, Joseph G.
;    APPLICANT:  Joshua O. Ojwang
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Oligonucleotides
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/145,704
;      FILING DATE:  28-OCT-1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/053,027
;      FILING DATE:  23-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5574-CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5151
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "Amine moity attached to 3'
;      OTHER INFORMATION:  end"
US-08-145-704-20
```

```
  Query Match           68.0%;  Score 13.6;  DB 1;  Length 26;
  Best Local Similarity  80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 5
US-08-145-704-25/c
; Sequence 25, Application US/08145704
; Patent No. 5567604
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Fennewald, Susan
;    APPLICANT:  Zendegui, Joseph G.
;    APPLICANT:  Joshua O. Ojwang
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Oligonucleotides
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/145,704
;      FILING DATE:  28-OCT-1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/053,027
;      FILING DATE:  23-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5574-CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5151
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "cholesterol moity attached
;      OTHER INFORMATION:  to 3' end"
US-08-145-704-25

  Query Match           68.0%;  Score 13.6;  DB 1;  Length 26;
  Best Local Similarity  80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 6
US-08-145-704-27/c
; Sequence 27, Application US/08145704
; Patent No. 5567604
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Fennewald, Susan
;    APPLICANT:  Zendegui, Joseph G.
;    APPLICANT:  Joshua O. Ojwang
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Oligonucleotides
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/145,704
;      FILING DATE:  28-OCT-1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/053,027
;      FILING DATE:  23-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5574-CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5151
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "phosphorothioate backbone
;      OTHER INFORMATION:  and amine moity attached to backbone"
US-08-145-704-27

  Query Match           68.0%;  Score 13.6;  DB 1;  Length 26;
  Best Local Similarity  80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 7
US-08-987-574-20/c
; Sequence 20, Application US/08987574
; Patent No. 6150339
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Fennewald, Susan
;    APPLICANT:  Zendegui, Joseph G.
;    APPLICANT:  Ojwang, Joshua O.
;    APPLICANT:  Hogan, Michael E.
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Oligonucleotides
;    NUMBER OF SEQUENCES:  52
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/987,574
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04529
;      FILING DATE:  28-OCT-1993
;      APPLICATION NUMBER:  US 08/053,027
;      FILING DATE:  23-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5574-CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5151
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "Amine moity
;      OTHER INFORMATION:  attached to 3' end"
US-08-987-574-20

  Query Match             68.0%;  Score 13.6;  DB 3;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           | |||||||  ||| |||||
Db      23 CCACCCCACCAACAACCCCC 4


RESULT 8
US-08-987-574-25/c
; Sequence 25, Application US/08987574
; Patent No. 6150339
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Fennewald, Susan
;    APPLICANT:  Zendegui, Joseph G.
;    APPLICANT:  Ojwang, Joshua O.
;    APPLICANT:  Hogan, Michael E.
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Oligonucleotides
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/987,574
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04529
;      FILING DATE:  28-OCT-1993
;      APPLICATION NUMBER:  US 08/053,027
;      FILING DATE:  23-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5574-CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5151
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "cholesterol moity
;      OTHER INFORMATION:  attached to 3' end"
US-08-987-574-25

  Query Match             68.0%;  Score 13.6;  DB 3;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           | |||||||  ||| |||||
Db      23 CCACCCCACCAACAACCCCC 4


RESULT 9
US-08-987-574-27/c
; Sequence 27, Application US/08987574
; Patent No. 6150339
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Fennewald, Susan
;    APPLICANT:  Zendegui, Joseph G.
;    APPLICANT:  Ojwang, Joshua O.
;    APPLICANT:  Hogan, Michael E.
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Oligonucleotides
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/987,574
;      FILING DATE:
;      CLASSIFICATION:  514
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04529
;      FILING DATE:  28-OCT-1993
;      APPLICATION NUMBER:  US 08/053,027
;      FILING DATE:  23-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5574-CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5151
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "phosphorothioate
;      OTHER INFORMATION:  backbone and amine moity attached to
;      OTHER INFORMATION:  backbone"
US-08-987-574-27

  Query Match             68.0%;  Score 13.6;  DB 3;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 10
US-08-535-168-20/c
; Sequence 20, Application US/08535168
; Patent No. 6184369
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Fennewald, Susan
;    APPLICANT:  Zendegui, Joseph G.
;    APPLICANT:  Ojwang, Joshua O.
;    APPLICANT:  Hogan, Michael E.
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Oligonucleotides
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/535,168
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04529
;      FILING DATE:  28-OCT-1993
;      APPLICATION NUMBER:  US 08/053,027
;      FILING DATE:  23-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5574-CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5151
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "Amine moity
;      OTHER INFORMATION:  attached to 3' end"
US-08-535-168-20

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 11
US-08-535-168-25/c
; Sequence 25, Application US/08535168
; Patent No. 6184369
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Fennewald, Susan
;    APPLICANT:  Zendegui, Joseph G.
;    APPLICANT:  Ojwang, Joshua O.
;    APPLICANT:  Hogan, Michael E.
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Oligonucleotides
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/535,168
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04529
;      FILING DATE:  28-OCT-1993
;      APPLICATION NUMBER:  US 08/053,027
;      FILING DATE:  23-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5574-CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5151
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "cholesterol moity
;      OTHER INFORMATION:  attached to 3' end"
US-08-535-168-25

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 12
US-08-535-168-27/c
; Sequence 27, Application US/08535168
; Patent No. 6184369
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Fennewald, Susan
;    APPLICANT:  Zendegui, Joseph G.
;    APPLICANT:  Ojwang, Joshua O.
;    APPLICANT:  Hogan, Michael E.
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Oligonucleotides
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/535,168
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/04529
;      FILING DATE:  28-OCT-1993
;      APPLICATION NUMBER:  US 08/053,027
;      FILING DATE:  23-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
;      REFERENCE/DOCKET NUMBER:  D-5574-CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/651-5151
;      TELEFAX:  713/651-5246
;      TELEX:  762829
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "phosphorothioate
;      OTHER INFORMATION:  backbone and amine moity attached to
;      OTHER INFORMATION:  backbone"
US-08-535-168-27

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 13
US-09-017-974-20/c
; Sequence 20, Application US/09017974
; Patent No. 6288042
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Ojwang, Joshua O.
;    APPLICANT:  Hogan, Michael E.
;    APPLICANT:  Wallace, Thomas L.
;    APPLICANT:  Cossum, Paul A.
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Tetrad Forming Oligonucleotides
;    NUMBER OF SEQUENCES:  88
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Conley, Rose & Tayon, P.C.
;      STREET:  600 Travis, Suite 1800
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77002-2912
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MS Word 97 (saved as .txt file)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/017,974
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,374
;      FILING DATE:  04-FEB-97
;      APPLICATION NUMBER:
;      FILING DATE:  09-DEC-97
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDaniel, C. Steven
;      REGISTRATION NUMBER:  33,962
;      REFERENCE/DOCKET NUMBER:  1472-06223
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/238-8010
;      TELEFAX:  713/238-8008
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "Amine moiety
;      OTHER INFORMATION:  attached to 3' end"
US-09-017-974-20

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 14
US-09-017-974-25/c
; Sequence 25, Application US/09017974
; Patent No. 6288042
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Ojwang, Joshua O.
;    APPLICANT:  Hogan, Michael E.
;    APPLICANT:  Wallace, Thomas L.
;    APPLICANT:  Cossum, Paul A.
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Tetrad Forming Oligonucleotides
;    NUMBER OF SEQUENCES:  88
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Conley, Rose & Tayon, P.C.
;      STREET:  600 Travis, Suite 1800
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77002-2912
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MS Word 97 (saved as .txt file)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/017,974
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,374
;      FILING DATE:  04-FEB-97
;      APPLICATION NUMBER:
;      FILING DATE:  09-DEC-97
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDaniel, C. Steven
;      REGISTRATION NUMBER:  33,962
;      REFERENCE/DOCKET NUMBER:  1472-06223
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/238-8010
;      TELEFAX:  713/238-8008
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "cholesterol moiety
;      OTHER INFORMATION:  attached to 3' end"
US-09-017-974-25

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 15
US-09-017-974-27/c
; Sequence 27, Application US/09017974
; Patent No. 6288042
;  GENERAL INFORMATION:
;    APPLICANT:  Rando, Robert F.
;    APPLICANT:  Ojwang, Joshua O.
;    APPLICANT:  Hogan, Michael E.
;    APPLICANT:  Wallace, Thomas L.
;    APPLICANT:  Cossum, Paul A.
;    TITLE OF INVENTION:  Anti-Viral Guanosine-Rich
;    TITLE OF INVENTION:  Tetrad Forming Oligonucleotides
;    NUMBER OF SEQUENCES:  88
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Conley, Rose & Tayon, P.C.
;      STREET:  600 Travis, Suite 1800
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77002-2912
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MS Word 97 (saved as .txt file)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/017,974
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,374
;      FILING DATE:  04-FEB-97
;      APPLICATION NUMBER:
;      FILING DATE:  09-DEC-97
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDaniel, C. Steven
;      REGISTRATION NUMBER:  33,962
;      REFERENCE/DOCKET NUMBER:  1472-06223
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713/238-8010
;      TELEFAX:  713/238-8008
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  26
;      OTHER INFORMATION:  /note= "phosphorothioate
;      OTHER INFORMATION:  backbone and amine moiety attached to
;      OTHER INFORMATION:  backbone"
US-09-017-974-27

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 8.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


Search completed: November 23, 2002, 06:36:13
Job time : 21.55 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-13
Perfect score: 20
Sequence: 1 tctctgtcctccaaatcgaa 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	14.8	74.0	71	3	PTU27511	U27511 Paramecium
c	2	14.2	71.0	84	6	A26456	A26456 Soya beta-1
c	3	14.2	71.0	84	6	I16749	I16749 Sequence 10
c	4	14.2	71.0	100	14	S68158S1	S68158 NS5 [5' and
	5	13.8	69.0	45	6	AR009531	AR009531 Sequence
	6	13.8	69.0	45	6	AR172143	AR172143 Sequence
	7	13.8	69.0	45	6	I66205	I66205 Sequence 12
	8	13.8	69.0	45	6	I66218	I66218 Sequence 12
	9	13.8	69.0	51	6	AX157723	AX157723 Sequence
	10	13.8	69.0	51	6	AX157724	AX157724 Sequence
	11	13.8	69.0	51	6	AX157725	AX157725 Sequence
	12	13.8	69.0	51	6	AX157726	AX157726 Sequence
c	13	13.6	68.0	30	6	AR118775	AR118775 Sequence
c	14	13.6	68.0	30	6	E51958	E51958 Diagnostic
	15	13.6	68.0	52	6	AR145888	AR145888 Sequence
	16	13.6	68.0	52	6	AX003018	AX003018 Sequence
	17	13.2	66.0	20	6	AX462752	AX462752 Sequence
c	18	13.2	66.0	45	6	AX195434	AX195434 Sequence
	19	13	65.0	16	6	AR021087	AR021087 Sequence
	20	13	65.0	16	6	AR043502	AR043502 Sequence
	21	13	65.0	16	6	AR062417	AR062417 Sequence
	22	13	65.0	16	6	AR183876	AR183876 Sequence
	23	13	65.0	17	6	AR190496	AR190496 Sequence
	24	13	65.0	18	6	AR190778	AR190778 Sequence
	25	12.8	64.0	29	6	AX297656	AX297656 Sequence
	26	12.8	64.0	30	6	AX297651	AX297651 Sequence
	27	12.8	64.0	30	6	AX297654	AX297654 Sequence
c	28	12.8	64.0	37	6	AR125431	AR125431 Sequence
	29	12.8	64.0	37	6	AR125433	AR125433 Sequence
	30	12.8	64.0	41	6	AR198452	AR198452 Sequence
c	31	12.8	64.0	48	6	AX426324	AX426324 Sequence
	32	12.8	64.0	51	6	AX204033	AX204033 Sequence
	33	12.8	64.0	51	6	AX297652	AX297652 Sequence
	34	12.8	64.0	51	6	AX297655	AX297655 Sequence
	35	12.8	64.0	52	6	AX297658	AX297658 Sequence
c	36	12.8	64.0	53	6	AR172220	AR172220 Sequence
	37	12.6	63.0	31	11	AL807511	AL807511 Arabidops
	38	12.6	63.0	75	8	YSCTGWA	M25440 Yeast tRNA-
	39	12.6	63.0	75	8	YSCTRW	M35060 Saccharomyc
c	40	12.6	63.0	81	6	I40761	I40761 Sequence 92
	41	12.6	63.0	89	17	HSEXTR35	Y10934 H.sapiens D
c	42	12.4	62.0	20	6	AR086399	AR086399 Sequence
c	43	12.4	62.0	20	6	AR122386	AR122386 Sequence
c	44	12.4	62.0	20	6	AR144251	AR144251 Sequence
c	45	12.4	62.0	20	6	AR209933	AR209933 Sequence

ALIGNMENTS

RESULT 1
PTU27511/c
LOCUS PTU27511 71 bp DNA linear INV 22-MAR-1996
DEFINITION Paramecium tetraurelia trichocyst matrix protein (T2c) gene, partial cds.
ACCESSION U27511
VERSION U27511.1 GI:881411
KEYWORDS .
SOURCE Paramecium tetraurelia.
ORGANISM Paramecium tetraurelia
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida; Paramecium.
REFERENCE 1 (bases 1 to 71)
AUTHORS Madeddu,L., Gautier,M.C., Le Caer,J.P., Garreau de Loubresse,N. and Sperling,L.

```
  TITLE     Protein processing and morphogenesis of secretory granules in
            Paramecium
  JOURNAL   Biochimie 76 (3-4), 329-335 (1994)
  MEDLINE   95119139
   PUBMED   7819344
REFERENCE   2  (bases 1 to 71)
  AUTHORS   Madeddu,L., Gautier,M.C., Vayssie,L., Houari,A. and Sperling,L.
  TITLE     A large multigene family codes for the polypeptides of the
            crystalline trichocyst matrix in Paramecium
  JOURNAL   Mol. Biol. Cell 6 (6), 649-659 (1995)
  MEDLINE   96059477
   PUBMED   7579685
REFERENCE   3  (bases 1 to 71)
  AUTHORS   Madeddu,L., Gautier,M.C. and Sperling,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-1995) Linda Sperling, Centre de Genetique
            Moleculaire, CNRS, Avenue de la Terasse, Gif-sur-Yvette 91198,
            France
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Paramecium tetraurelia"
                     /strain="d4-2"
                     /db_xref="taxon:5888"
     gene            1. .71
                     /gene="T2c"
     CDS             <1. .>71
                     /gene="T2c"
                     /note="member of the Paramecium trichocyst multigene
                     family; Paramecium secretory granule protein"
                     /codon_start=1
                     /transl_table=6
                     /product="trichocyst matrix protein T2"
                     /protein_id="AAA92611.1"
                     /db_xref="GI:881412"
                     /translation="DPLDRLLQTLTDLEDRYVAEQKE"
BASE COUNT       26 a     12 c     15 g     18 t
ORIGIN

  Query Match           74.0%;  Score 14.8;  DB 3;  Length 71;
  Best Local Similarity  88.9%;  Pred. No. 9.9e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATCG 18
           | |||||| |||||||||
Db      50 TATCTGTCTTCCAAATCG 33


RESULT 2
A26456/c
LOCUS       A26456                   84 bp    DNA     linear   PAT 22-SEP-1995
DEFINITION  Soya beta-1,3-glucanase C-terminal sequence.
ACCESSION   A26456
VERSION     A26456.1  GI:1247332
KEYWORDS    .
SOURCE      Glycine max.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
BASE COUNT       25 a     12 c     23 g     24 t
ORIGIN

  Query Match           71.0%;  Score 14.2;  DB 6;  Length 84;
  Best Local Similarity  84.2%;  Pred. No. 2.4e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CTCTGTCCTCCAAATCGAA 20
           |||| ||||||||| | ||
Db      26 CTCTTTCCTCCAAACCCAA 8


RESULT 3
I16749/c
LOCUS       I16749                   84 bp    DNA     linear   PAT 03-APR-1996
DEFINITION  Sequence 10 from patent US 5477001.
ACCESSION   I16749
VERSION     I16749.1  GI:1251657
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 84)
  AUTHORS   Sass,C., Leguay,J.-J., Grison,R. and Toppan,A.
  TITLE     Recombinant DNA coding for a novel protein having
            .beta.-1,3-glucanase activity, bacteria containing this DNA,
            transformed plant cells and plants
  JOURNAL   Patent: US 5477001-A 10 19-DEC-1995;
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="unknown"
BASE COUNT       25 a     12 c     23 g     24 t
ORIGIN

  Query Match           71.0%;  Score 14.2;  DB 6;  Length 84;
  Best Local Similarity  84.2%;  Pred. No. 2.4e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CTCTGTCCTCCAAATCGAA 20
           |||| ||||||||| | ||
Db      26 CTCTTTCCTCCAAACCCAA 8


RESULT 4
S68158S1/c
LOCUS       S68158S1                 100 bp    RNA     linear   VRL 23-SEP-1994
DEFINITION  NS5 {5' and 3' conserved regions} [central european tick-borne
            encephalitis virus Kumlinge, Genomic RNA, 100 nt, segment 1 of 2].
ACCESSION   S68158
VERSION     S68158.1  GI:544621
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Kumlinge virus.
  ORGANISM  Kumlinge virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Flavivirus; tick-borne encephalitis virus group.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Fulop,L., Barrett,A.D., Phillpotts,R., Martin,K., Leslie,D. and
            Titball,R.W.
  TITLE     Rapid identification of flaviviruses based on conserved NS5 gene
            sequences
  JOURNAL   J. Virol. Methods 44 (2-3), 179-188 (1993)
  MEDLINE   94086710
   PUBMED   8263114
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 142831] from the original journal article.
            This sequence comes from Fig. 3a and 3b.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Kumlinge virus"
                     /db_xref="taxon:11092"
BASE COUNT       26 a     26 c     29 g     19 t
ORIGIN

  Query Match           71.0%;  Score 14.2;  DB 14;  Length 100;
  Best Local Similarity  84.2%;  Pred. No. 2.5e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATCGA 19
           ||||||  ||||||| |||
Db      83 TCTCTGGTCTCCAAACCGA 65
```

```
RESULT 5
AR009531
LOCUS       AR009531                 45 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 12 from patent US 5756312.
ACCESSION   AR009531
VERSION     AR009531.1  GI:3968336
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Weiner,A.J. and Houghton,M.
  TITLE     Immunoreactive polypeptide compositions
  JOURNAL   Patent: US 5756312-A 12 26-MAY-1998;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT        8 a     16 c      6 g     15 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 45;
  Best Local Similarity   88.2%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATC 17
          || |||||||||| ||
Db     29 TCCCTGTCCTCCAAGTC 45


RESULT 6
AR172143
LOCUS       AR172143                 45 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 12 from patent US 6303292.
ACCESSION   AR172143
VERSION     AR172143.1  GI:17911634
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Weiner,A.J. and Houghton,M.
  TITLE     Immunoreactive polypeptide compositions
  JOURNAL   Patent: US 6303292-A 12 16-OCT-2001;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT        8 a     16 c      6 g     15 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 45;
  Best Local Similarity   88.2%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATC 17
          || |||||||||| ||
Db     29 TCCCTGTCCTCCAAGTC 45


RESULT 7
I66205
LOCUS       I66205                   45 bp    DNA     linear   PAT 28-DEC-1997
DEFINITION  Sequence 12 from patent US 5670152.
ACCESSION   I66205
VERSION     I66205.1  GI:2724182
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Weiner,A.J. and Houghton,M.
  TITLE     Immunoreactive polypeptide compositions
  JOURNAL   Patent: US 5670152-A 12 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT        8 a     16 c      6 g     15 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 45;
  Best Local Similarity   88.2%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATC 17
          || |||||||||| ||
Db     29 TCCCTGTCCTCCAAGTC 45


RESULT 8
I66218
LOCUS       I66218                   45 bp    DNA     linear   PAT 28-DEC-1997
DEFINITION  Sequence 12 from patent US 5670153.
ACCESSION   I66218
VERSION     I66218.1  GI:2724195
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Weiner,A.J. and Houghton,M.
  TITLE     Immunoreactive polypeptide compositions
  JOURNAL   Patent: US 5670153-A 12 23-SEP-1997;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT        8 a     16 c      6 g     15 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 45;
  Best Local Similarity   88.2%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATC 17
          || |||||||||| ||
Db     29 TCCCTGTCCTCCAAGTC 45


RESULT 9
AX157723
LOCUS       AX157723                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 1051 from Patent WO0140521.
ACCESSION   AX157723
VERSION     AX157723.1  GI:14539054
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 1051 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (1052 is other entry)
                     Accession number cg27807001"
BASE COUNT       11 a     11 c     19 g     10 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 3.7e+03;
```

```
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CTGTCCTCCAAATCGAA 20
          |||||||| || |||||
Db      7 CTGTCCTCGAAGTCGAA 23


RESULT 10
AX157724
LOCUS       AX157724                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 1052 from Patent WO0140521.
ACCESSION   AX157724
VERSION     AX157724.1  GI:14539055
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 1052 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (1051 is other entry)
                     Accession number cg27807001"
BASE COUNT       12 a     11 c     18 g     10 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 3.7e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CTGTCCTCCAAATCGAA 20
          |||||||| || |||||
Db      7 CTGTCCTCGAAGTCGAA 23


RESULT 11
AX157725
LOCUS       AX157725                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 1053 from Patent WO0140521.
ACCESSION   AX157725
VERSION     AX157725.1  GI:14539056
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 1053 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (1054 is other entry)
                     Accession number cg27807001"
BASE COUNT       11 a     12 c     20 g      8 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 3.7e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CTGTCCTCCAAATCGAA 20
          |||||||| || |||||
Db      1 CTGTCCTCGAAGTCGAA 17


RESULT 12
AX157726
LOCUS       AX157726                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 1054 from Patent WO0140521.
ACCESSION   AX157726
VERSION     AX157726.1  GI:14539057
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 1054 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (1053 is other entry)
                     Accession number cg27807001"
BASE COUNT       10 a     12 c     21 g      8 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 3.7e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CTGTCCTCCAAATCGAA 20
          |||||||| || |||||
Db      1 CTGTCCTCGAAGTCGAA 17


RESULT 13
AR118775/c
LOCUS       AR118775                 30 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 205 from patent US 6150087.
ACCESSION   AR118775
VERSION     AR118775.1  GI:14100685
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Chien,D.Y.
  TITLE     NANBV diagnostics and vaccines
  JOURNAL   Patent: US 6150087-A 205 21-NOV-2000;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        9 a      5 c     12 g      4 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 4.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATCGAA 20
          || ||||| |||| |||| |
Db     26 TCCCTGTCTTCCAGATCGCA 7


RESULT 14
E51958/c
```

```
LOCUS       E51958                    30 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Diagnostic for NANBV.
ACCESSION   E51958
VERSION     E51958.1  GI:13019285
KEYWORDS    JP 2000039434-A/1.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Michael,H., Quilim,C. and George,K.
  TITLE     Diagnostic for NANBV
  JOURNAL   Patent: JP 2000039434-A 1 08-FEB-2000;
            CHIRON CORP
COMMENT     OS   Unidentified
            PN   JP 2000039434-A/1
            PD   08-FEB-2000
            PF   03-JUN-1999 JP 1999157192
            PR   17-MAR-1989 US      325.338,20-APR-1989 US      341.334, PR
            18-MAY-1989 US      355.002
            PI   MICHAEL HOUGHTON,QUI-LIM CHOO,GEORGE KUO
            PC   G01N33/569,C07K14/18,G01N33/576//C12N15/09,C12P21/02, PC
            (C12P21/02,C12R1:19),
            PC   C12N15/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .30
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        9 a      5 c     12 g      4 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 30;
  Best Local Similarity  80.0%;  Pred. No. 4.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATCGAA 20
           || ||||| |||| |||| |
Db      26 TCCCTGTCTTCCAGATCGCA 7


RESULT 15
AR145888
LOCUS       AR145888                  52 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 8 from patent US 6218139.
ACCESSION   AR145888
VERSION     AR145888.1  GI:15109077
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Achstetter,T., Cauet,G. and Degryse,E.
  TITLE     Yeast strains possessing the interrupted ATF2 gene and their
            applications
  JOURNAL   Patent: US 6218139-A 8 17-APR-2001;
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="unknown"
BASE COUNT       20 a     12 c      9 g     11 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 52;
  Best Local Similarity  80.0%;  Pred. No. 4.9e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATCGAA 20
           ||||||||  | |||| |||
Db       2 TCTCTGTCGACAAAATGGAA 21
```

Search completed: December 3, 2002, 18:13:24
Job time : 309.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

QM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-13
Perfect score: 20
Sequence: 1 tctctgtcctccaaatcgaa 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14	70.0	63	9	US-09-933-797-384	Sequence 384, App
	2	13.6	68.0	52	10	US-09-797-239-8	Sequence 8, Appli
c	3	13.2	66.0	45	10	US-09-760-139-33	Sequence 33, Appl
c	4	12.4	62.0	26	9	US-09-905-291A-381	Sequence 381, App
c	5	12.4	62.0	26	10	US-09-909-320-381	Sequence 381, App
c	6	12.4	62.0	26	10	US-09-909-088B-381	Sequence 381, App
c	7	12.4	62.0	95	10	US-09-969-373-1290	Sequence 1290, Ap
c	8	12.4	62.0	97	10	US-09-969-373-838	Sequence 838, App
	9	12.2	61.0	30	9	US-09-733-042-2	Sequence 2, Appli
c	10	12.2	61.0	30	12	US-10-041-400A-5	Sequence 5, Appli
c	11	12.2	61.0	30	12	US-10-041-264A-5	Sequence 5, Appli
c	12	12.2	61.0	30	12	US-10-042-091A-5	Sequence 5, Appli
c	13	12.2	61.0	33	9	US-10-040-803-10	Sequence 10, Appl
	14	12.2	61.0	67	10	US-09-815-242-1629	Sequence 1629, Ap
	15	12.2	61.0	96	10	US-09-828-538-17	Sequence 17, Appl
	16	12.2	61.0	96	10	US-09-878-574-5985	Sequence 5985, Ap
c	17	12	60.0	68	10	US-09-760-139-36	Sequence 36, Appl
c	18	11.8	59.0	23	10	US-09-950-902-13	Sequence 13, Appl
c	19	11.8	59.0	31	10	US-09-801-274-1171	Sequence 1171, Ap
c	20	11.8	59.0	86	10	US-09-864-761-31474	Sequence 31474, A
c	21	11.8	59.0	87	10	US-09-878-574-13597	Sequence 13597, A
c	22	11.8	59.0	100	10	US-09-864-761-19163	Sequence 19163, A
	23	11.8	59.0	100	10	US-09-864-761-33290	Sequence 33290, A
c	24	11.6	58.0	31	8	US-08-834-666A-59	Sequence 59, Appl
c	25	11.6	58.0	48	10	US-09-565-233-7	Sequence 7, Appli
	26	11.6	58.0	52	10	US-09-962-628B-41	Sequence 41, Appl
	27	11.6	58.0	52	10	US-09-962-628B-42	Sequence 42, Appl
c	28	11.6	58.0	98	10	US-09-878-574-14881	Sequence 14881, A
	29	11.4	57.0	18	10	US-09-969-373-1572	Sequence 1572, Ap
	30	11.4	57.0	19	10	US-09-832-424-10	Sequence 10, Appl
c	31	11.4	57.0	31	10	US-09-801-274-259	Sequence 259, App
c	32	11.4	57.0	87	10	US-09-834-975-62	Sequence 62, Appl
	33	11.4	57.0	95	10	US-09-878-574-920	Sequence 920, App
	34	11.4	57.0	100	10	US-09-969-373-1195	Sequence 1195, Ap
c	35	11.2	56.0	16	10	US-09-823-847-38	Sequence 38, Appl
	36	11.2	56.0	18	10	US-09-969-373-3820	Sequence 3820, Ap
	37	11.2	56.0	20	10	US-09-735-995-62	Sequence 62, Appl
	38	11.2	56.0	21	10	US-09-844-493-16	Sequence 16, Appl
	39	11.2	56.0	21	10	US-09-844-501-16	Sequence 16, Appl
	40	11.2	56.0	21	10	US-09-844-265-16	Sequence 16, Appl
	41	11.2	56.0	23	10	US-09-759-667A-9	Sequence 9, Appli
c	42	11.2	56.0	28	10	US-09-898-323-7	Sequence 7, Appli
	43	11.2	56.0	28	10	US-09-898-323-16	Sequence 16, Appl
c	44	11.2	56.0	29	10	US-09-957-974-10	Sequence 10, Appl
	45	11.2	56.0	29	10	US-09-898-323-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-933-797-384
; Sequence 384, Application US/09933797
; Patent No. US20020155119A1
; GENERAL INFORMATION:
; APPLICANT: Robert A. Sikes et al.
; TITLE OF INVENTION: Isolation and Use of Fetal Urogenital
; TITLE OF INVENTION: Sinus Expressed Sequences
; FILE REFERENCE: 9901-007-999
; CURRENT APPLICATION NUMBER: US/09/933,797
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US/09/482,933
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: PCT/US99/10746
; PRIOR FILING DATE: 1999-05/14
; PRIOR APPLICATION NUMBER: 60/085,383
; PRIOR FILING DATE: 1998-05-14
; NUMBER OF SEQ ID NOS: 811
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 384
; LENGTH: 63
; TYPE: DNA
; ORGANISM: Murine
US-09-933-797-384

Query Match 70.0%; Score 14; DB 9; Length 63;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TCTGTCCTCCAAAT 16
||||||||||||||
Db 50 TCTGTCCTCCAAAT 63

RESULT 2
US-09-797-239-8
; Sequence 8, Application US/09797239
; Patent No. US20010012630A1
; GENERAL INFORMATION:
; APPLICANT: Hoechst Marion Roussel
; TITLE OF INVENTION: Yeast strains possessing the interrupted ATF2 gene and
; TITLE OF INVENTION: their applications.

```
;  FILE REFERENCE:  2482 PCT SEQUENCES IN ENGLISH
;  CURRENT APPLICATION NUMBER: US/09/797,239
;  CURRENT FILING DATE:  2001-03-01
;  NUMBER OF SEQ ID NOS:  17
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 52
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: OLIGONUCLEOTIDE
US-09-797-239-8

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 52;
  Best Local Similarity   80.0%;  Pred. No. 5.5e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATCGAA 20
           ||||||||  | ||||  |||
Db       2 TCTCTGTCGACAAAATGGAA 21


RESULT 3
US-09-760-139-33/c
; Sequence 33, Application US/09760139
; Patent No. US20020058304A1
; GENERAL INFORMATION:
;  APPLICANT:  Yaver, Debbie S.
;  APPLICANT:  Bellini, Daniel A.
;  TITLE OF INVENTION:  Methods For Producing A Polypeptide
;  TITLE OF INVENTION:  Using A Crippled Translational Initiator Sequence
;  FILE REFERENCE:  5966.200-US
;  CURRENT APPLICATION NUMBER: US/09/760,139
;  CURRENT FILING DATE:  2001-01-12
;  PRIOR APPLICATION NUMBER:  09/482,751
;  PRIOR FILING DATE:  2000-01-13
;  NUMBER OF SEQ ID NOS:  36
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 33
;   LENGTH: 45
;   TYPE: DNA
;   ORGANISM: Aspergillus oryzae
US-09-760-139-33

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 45;
  Best Local Similarity   83.3%;  Pred. No. 8.4e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATCG 18
           | | ||||||||||  |||
Db      28 TTTATGTCCTCCAAGTCG 11


RESULT 4
US-09-905-291A-381/c
; Sequence 381, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/905,291A
;  CURRENT FILING DATE:  2001-07-12
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 381
;   LENGTH: 26
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic oligonucleotide probe
US-09-905-291A-381

  Query Match             62.0%;  Score 12.4;  DB 9;  Length 26;
  Best Local Similarity   92.9%;  Pred. No. 1.9e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 CTGTCCTCCAAATC 17
           |||||||| |||||
Db      17 CTGTCCTCAAAATC 4


RESULT 5
US-09-909-320-381/c
; Sequence 381, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
```

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 381
; LENGTH: 26
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide probe
US-09-909-320-381

Query Match 62.0%; Score 12.4; DB 10; Length 26;
Best Local Similarity 92.9%; Pred. No. 1.9e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CTGTCCTCCAAATC 17
||||||||.|||||
Db 17 CTGTCCTCAAAATC 4

RESULT 6
US-09-909-088B-381/c
; Sequence 381, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20

```
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 381
;   LENGTH: 26
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic oligonucleotide probe
US-09-909-088B-381

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 26;
  Best Local Similarity   92.9%;  Pred. No. 1.9e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 CTGTCCTCCAAATC 17
           |||||||| |||||
Db      17 CTGTCCTCAAAATC 4


RESULT 7
US-09-969-373-1290/c
; Sequence 1290, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
;  APPLICANT: Effertz, Roger J.
;  APPLICANT:  Hauge, Brian M.
;  TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
;  FILE REFERENCE: 38-10(52679)A
;  CURRENT APPLICATION NUMBER: US/09/969,373
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US 09/754,853
;  PRIOR FILING DATE: 2001-01-05
;  PRIOR APPLICATION NUMBER: US 09/760,427
;  PRIOR FILING DATE: 2001-01-13
;  PRIOR APPLICATION NUMBER: US 09/855,768
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 1290
;   LENGTH: 95
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-969-373-1290

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 95;
  Best Local Similarity   92.9%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 TCCTCCAAATCGAA 20
           | ||||||||||||
Db      73 TGCTCCAAATCGAA 60


RESULT 8
US-09-969-373-838/c
; Sequence 838, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
;  APPLICANT: Effertz, Roger J.
;  APPLICANT:  Hauge, Brian M.
;  TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
;  FILE REFERENCE: 38-10(52679)A
;  CURRENT APPLICATION NUMBER: US/09/969,373
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US 09/754,853
;  PRIOR FILING DATE: 2001-01-05
;  PRIOR APPLICATION NUMBER: US 09/760,427
;  PRIOR FILING DATE: 2001-01-13
;  PRIOR APPLICATION NUMBER: US 09/855,768
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 838
;   LENGTH: 97
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-969-373-838

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 97;
  Best Local Similarity   92.9%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAA 14
           ||||| ||||||||
Db      67 TCTCTCTCCTCCAA 54


RESULT 9
US-09-733-042-2
; Sequence 2, Application US/09733042
; Patent No. US20020168709A1
; GENERAL INFORMATION:
;  APPLICANT:   Hennecke, Frank
;  APPLICANT:  Renner, Wolfgang A.
;  TITLE OF INVENTION:   Replicon Based Activation of Endogenous Genes
;  FILE REFERENCE:   1700.0100001
;  CURRENT APPLICATION NUMBER: US/09/733,042
;  CURRENT FILING DATE:  2000-12-11
;  PRIOR APPLICATION NUMBER:   US 60/169,988
;  PRIOR FILING DATE:   1999-12-10
;  NUMBER OF SEQ ID NOS:   49
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Enh-FOR
US-09-733-042-2

  Query Match             61.0%;  Score 12.2;  DB 9;  Length 30;
  Best Local Similarity   82.4%;  Pred. No. 2.5e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATC 17
           ||||   ||||||||||
Db       5 TCTCGAGCCTCCAAATC 21


RESULT 10
US-10-041-400A-5/c
; Sequence 5, Application US/10041400A
; Patent No. US20020110895A1
; GENERAL INFORMATION:
;  APPLICANT: Darrow, Andrew
;  APPLICANT:  Andrade-Gordon, Patricia
;  APPLICANT:  Qi, Jenson
;  TITLE OF INVENTION: DNA Encoding the Human Serine
;  TITLE OF INVENTION:  Protease EOS
;  FILE REFERENCE: ORT-1031
;  CURRENT APPLICATION NUMBER: US/10/041,400A
;  CURRENT FILING DATE:  2002-01-08
;  PRIOR APPLICATION NUMBER: US/09/387,375
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: oligonucleotide
US-10-041-400A-5

  Query Match             61.0%;  Score 12.2;  DB 12;  Length 30;
  Best Local Similarity   82.4%;  Pred. No. 2.5e+03;
```

```
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATC 17
          |||| |||||| | |||
Db     19 TCTCCGTCCTCTAGATC 3


RESULT 11
US-10-041-264A-5/c
; Sequence 5, Application US/10041264A
; Patent No. US20020142446A1
; GENERAL INFORMATION:
;  APPLICANT: Darrow, Andrew
;  APPLICANT:  Andrade-Gordon, Patricia
;  APPLICANT:  Qi, Jenson
;  TITLE OF INVENTION: DNA Encoding the Human Serine
;  TITLE OF INVENTION:  Protease EOS
;  FILE REFERENCE: ORT-1031
;  CURRENT APPLICATION NUMBER: US/10/041,264A
;  CURRENT FILING DATE:  2002-01-08
;  PRIOR APPLICATION NUMBER: US/09/387,375
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: oligonucleotide
US-10-041-264A-5

  Query Match            61.0%;  Score 12.2;  DB 12;  Length 30;
  Best Local Similarity   82.4%;  Pred. No. 2.5e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATC 17
          |||| |||||| | |||
Db     19 TCTCCGTCCTCTAGATC 3


RESULT 12
US-10-042-091A-5/c
; Sequence 5, Application US/10042091A
; Patent No. US20020142447A1
; GENERAL INFORMATION:
;  APPLICANT: Darrow, Andrew
;  APPLICANT:  Andrade-Gordon, Patricia
;  APPLICANT:  Qi, Jenson
;  TITLE OF INVENTION: DNA Encoding the Human Serine
;  TITLE OF INVENTION:  Protease EOS
;  FILE REFERENCE: ORT-1031
;  CURRENT APPLICATION NUMBER: US/10/042,091A
;  CURRENT FILING DATE:  2002-01-08
;  PRIOR APPLICATION NUMBER: US/09/387,375
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: oligonucleotide
US-10-042-091A-5

  Query Match            61.0%;  Score 12.2;  DB 12;  Length 30;
  Best Local Similarity   82.4%;  Pred. No. 2.5e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATC 17
          |||| |||||| | |||
Db     19 TCTCCGTCCTCTAGATC 3


RESULT 13
US-10-040-803-10/c
; Sequence 10, Application US/10040803
; Patent No. US20020164767A1
; GENERAL INFORMATION:
;  APPLICANT: Andrade-Gordon, Patricia
;  APPLICANT:  Darrow, Andrew L.
;  APPLICANT:  Qi, Jenson
;  TITLE OF INVENTION: DNA encoding human serine protease C-E
;  FILE REFERENCE: ORT-1030
;  CURRENT APPLICATION NUMBER: US/10/040,803
;  CURRENT FILING DATE:  2002-01-07
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: primer
;   OTHER INFORMATION: oligonucleotide
US-10-040-803-10

  Query Match            61.0%;  Score 12.2;  DB 9;  Length 33;
  Best Local Similarity   82.4%;  Pred. No. 2.5e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATC 17
          || |||||||| | |||
Db     19 TCGCTGTCCTCTAGATC 3


RESULT 14
US-09-815-242-1629
; Sequence 1629, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1629
;   LENGTH: 67
;   TYPE: DNA
```

; ORGANISM: Staphylococcus aureus
US-09-815-242-1629

Query Match 61.0%; Score 12.2; DB 10; Length 67;
Best Local Similarity 82.4%; Pred. No. 2.8e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TCTCTGTCCTCCAAATC 17
||||| ||| ||||||
Db 3 TCTCTCTCCAGCAAATC 19

RESULT 15
US-09-828-538-17
; Sequence 17, Application US/09828538
; Patent No. US20010025031A1
; GENERAL INFORMATION:
; APPLICANT: Ellis, Catherine E.
; APPLICANT: Kwok, Cheni
; APPLICANT: Bodsworth, Nicola J.
; APPLICANT: Halsey, Wendy
; APPLICANT: Van Horn, Stephanie
; TITLE OF INVENTION: HFGAN72 Receptor Genomic DNA and Methods
; TITLE OF INVENTION: of Use Thereof in Diagnostic Applications
; FILE REFERENCE: GH-50038-C1
; CURRENT APPLICATION NUMBER: US/09/828,538
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/088,624
; PRIOR FILING DATE: 1998-06-08
; PRIOR APPLICATION NUMBER: 60/093,726
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: 09/328,014
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 96
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-828-538-17

Query Match 61.0%; Score 12.2; DB 10; Length 96;
Best Local Similarity 82.4%; Pred. No. 2.9e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTCTGTCCTCCAAATCG 18
||||||||| | ||||
Db 22 CTCTGTCCTGCCCATCG 38

Search completed: November 23, 2002, 06:42:10
Job time : 18.25 secs

```
          |||||   ||||||||||
Db      26 GCCGAACAGGGACCCAGA 9


RESULT 6
AR127237
LOCUS       AR127237                  33 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 41 from patent US 6180391.
ACCESSION   AR127237
VERSION     AR127237.1  GI:14113830
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Brown,W.C.
  TITLE     Highly efficient controlled expression of exogenous genes in e.
            coli
  JOURNAL   Patent: US 6180391-A 41 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="unknown"
BASE COUNT        4 a      11 c      11 g       7 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 33;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAG 19
           |||| ||||||| || ||
Db      12 TGCCTACGTGGGTCCTAG 29


RESULT 7
AR127238/c
LOCUS       AR127238                  33 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 42 from patent US 6180391.
ACCESSION   AR127238
VERSION     AR127238.1  GI:14113831
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Brown,W.C.
  TITLE     Highly efficient controlled expression of exogenous genes in e.
            coli
  JOURNAL   Patent: US 6180391-A 42 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="unknown"
BASE COUNT        8 a      10 c      10 g       5 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 33;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAG 19
           |||| ||||||| || ||
Db      26 TGCCTACGTGGGTCCTAG 9


RESULT 8
AX358674/c
LOCUS       AX358674                  35 bp    DNA     linear   PAT 13-FEB-2002
DEFINITION  Sequence 52 from Patent WO0190419.
ACCESSION   AX358674
VERSION     AX358674.1  GI:18675208
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Stanton,V.P.
  TITLE     Methods for genetic analysis of dna to detect sequence variances
  JOURNAL   Patent: WO 0190419-A 52 29-NOV-2001;
            Variagenics, Inc. (US) ; Stanton, Vincent P., Jr. (US)
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Exemplary motif"
BASE COUNT        9 a      11 c       6 g       8 t       1 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 35;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCCGACGTGGGACCCAGA 20
           ||||| |||| || ||||
Db      29 GCCGAAGTGGAACRCAGA 12


RESULT 9
AX358675
LOCUS       AX358675                  35 bp    DNA     linear   PAT 13-FEB-2002
DEFINITION  Sequence 53 from Patent WO0190419.
ACCESSION   AX358675
VERSION     AX358675.1  GI:18675209
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Stanton,V.P.
  TITLE     Methods for genetic analysis of dna to detect sequence variances
  JOURNAL   Patent: WO 0190419-A 53 29-NOV-2001;
            Variagenics, Inc. (US) ; Stanton, Vincent P., Jr. (US)
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Exemplary motif"
BASE COUNT        8 a       6 c      11 g       9 t       1 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 35;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCCGACGTGGGACCCAGA 20
           ||||| |||| || ||||
Db       7 GCCGAAGTGGAACRCAGA 24


RESULT 10
AR143578/c
LOCUS       AR143578                  44 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 34 from patent US 6204371.
ACCESSION   AR143578
VERSION     AR143578.1  GI:15104864
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Levinson,D.Adam.
  TITLE     Compositions and methods for the treatment and diagnosis of immune
            disorders
  JOURNAL   Patent: US 6204371-A 34 20-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="unknown"
```

```
BASE COUNT        9 a     15 c     10 g     10 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 44;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCCGACGTGGGACCCA 18
          |||| ||| ||||| |||
Db     33 GTGCAGACTTGGGATCCA 16


RESULT 11
AR168947/c
LOCUS       AR168947                 44 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 34 from patent US 6288218.
ACCESSION   AR168947
VERSION     AR168947.1  GI:17905145
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Levinson,D.Adam.
  TITLE     Compositions and methods for the treatment and diagnosis of immune
            disorders
  JOURNAL   Patent: US 6288218-A 34 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="unknown"
BASE COUNT        9 a     15 c     10 g     10 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 44;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCCGACGTGGGACCCA 18
          |||| ||| ||||| |||
Db     33 GTGCAGACTTGGGATCCA 16


RESULT 12
HSEXTR2
LOCUS       HSEXTR2                  90 bp    DNA     linear   PRI 15-APR-1997
DEFINITION  H.sapiens exon trapped from 6q27, 4J4VSex100c4.
ACCESSION   X99574
VERSION     X99574.1  GI:1490304
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   Barua,U., Rousseaux,S. and Edwards,Y.H.
  TITLE     Identification of exons in a region of human chromosome 6q known to
            contain tumour suppressor genes
  JOURNAL   DNA Seq. 7 (1), 13-19 (1996)
  MEDLINE   97217669
   PUBMED   9063630
REFERENCE   2  (bases 1 to 90)
  AUTHORS   Edwards,Y.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-1996) Y.H. Edwards, Wolfson House, MRC Human
            Biochemical Genetics Unit, UCL, 4 Stephenson Way, London, NW1 2HE,
            UK
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="q27"
                     /clone_lib="IFGT 4J4"
     gene            1. .90
                     /gene="4J4VSex11c4"
     exon            1. .90
                     /gene="4J4VSex11c4"
BASE COUNT       26 a     23 c     18 g     23 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 90;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGCCGACGTGGGACCCAG 19
          |||| ||||| ||| |||
Db     68 TGCCCACGTGTGACACAG 85


RESULT 13
E11187
LOCUS       E11187                   22 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Probe.
ACCESSION   E11187
VERSION     E11187.1  GI:22024828
KEYWORDS    JP 1996089300-A/2.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Tokuda,C., Fukuda,T. and Saito,I.
  TITLE     METHOD FOR DETECTING NUCLEIC ACID
  JOURNAL   Patent: JP 1996089300-A 2 09-APR-1996;
            MITSUI TOATSU CHEM INC, MITSUI PHARMACEUT INC
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1996089300-A/2
            PD   09-APR-1996
            PF   28-SEP-1994 JP 1994232637
            PI   TOKUDA CHIKASHI, FUKUDA TAMOTSU, SAITO IZUMI
            PC   C12Q1/68,G01N33/58//C12N15/09,C12Q1/70;
            CC   strandedness: Single;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .22
            FT                   /organism='Artificial sequences'.
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        7 a      6 c      7 g      2 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 22;
  Best Local Similarity   87.5%;  Pred. No. 5.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CCGACGTGGGACCCAG 19
          ||||  ||||||||||
Db      1 CCGAAATGGGACCCAG 16


RESULT 14
AX019309/c
LOCUS       AX019309                 24 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 9 from Patent WO9940221.
ACCESSION   AX019309
VERSION     AX019309.1  GI:10043296
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Dahm,M.W., Phelps,R.C. and Brockmeyer,C.
```

```
  TITLE     Method for quantitatively analyzing tumor cells in a body fluid and
            test kits suited therefor
  JOURNAL   Patent: WO 9940221-A 9 12-AUG-1999;
            DAHM MICHAEL W (DE); PHELPS ROBERT C (DE); BROCKMEYER CARSTEN (DE)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        3 a      9 c      6 g      6 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 24;
  Best Local Similarity   87.5%;  Pred. No. 5.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 CGACGTGGGACCCAGA 20
           || ||||||| |||||
Db      19 CGTCGTGGGAGCCAGA 4


RESULT 15
AR065111
LOCUS       AR065111                  25 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 1 from patent US 5849487.
ACCESSION   AR065111
VERSION     AR065111.1  GI:5995327
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Hase,T. and Masubuchi,H.
  TITLE     Nucleic acid detection by repetitive synthesis and degradation of a
            nucleotide on a nuclease resistant oligonucleotide
  JOURNAL   Patent: US 5849487-A 1 15-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="unknown"
BASE COUNT        7 a      8 c      7 g      3 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 5.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCGACGTGGGACCCAG 19
           ||||  ||||||||||
Db       3 CCGAAATGGGACCCAG 18


Search completed: December  3, 2002, 18:13:03
Job time : 309.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-10
Perfect score: 20
Sequence: 1 gtgccgacgtgggacccaga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31440	Human neuropilin m
c	2	18	90.0	21	22	AAF97678	Human gene single
	3	13.8	69.0	27	22	AAS07631	Streptococcus coel
c	4	13.8	69.0	28	20	AAX83785	PmeI restriction e
c	5	13.8	69.0	50	22	AAL34511	Human SNP oligonuc
	6	13.8	69.0	65	24	ABN29579	Rat spliced transc
	7	13.6	68.0	27	20	AAX23277	H. vulgare cystein
c	8	13.6	68.0	51	22	AAI79205	Human silent SNP c
	9	13.6	68.0	65	24	ABN29322	Rat spliced transc
c	10	13.6	68.0	65	24	ABN29817	Rat spliced transc
c	11	13.4	67.0	50	22	AAL34194	Human SNP oligonuc
c	12	13.2	66.0	27	20	AAX03341	PCR primer used to
	13	13.2	66.0	33	20	AAX89941	Sequence of primer
c	14	13.2	66.0	33	20	AAX89942	Sequence of primer
c	15	13.2	66.0	35	24	ABL32002	Dihydropyrimidine
c	16	13.2	66.0	44	17	AAT38282	Murine 103 gene 3'
c	17	13.2	66.0	44	21	AAA51915	Reverse primer to
c	18	13.2	66.0	44	22	AAI70278	Mouse 103 gene rev
c	19	13.2	66.0	44	22	AAD03383	3' primer, to ampl
c	20	13.2	66.0	44	22	AAF82626	Murine TH2-enriche
c	21	13.2	66.0	44	22	AAC92152	Mouse 103 gene PCR
c	22	13.2	66.0	44	22	AAF23475	3' oligonucleotide
	23	13.2	66.0	65	24	ABN31280	Rat spliced transc
c	24	13.2	66.0	83	21	AAC17161	Human secreted pro
	25	12.8	64.0	22	17	AAT36509	Synthetic probe A
c	26	12.8	64.0	24	20	AAZ00723	Human telomerase c
	27	12.8	64.0	25	16	AAQ92893	Nuclease resistant
c	28	12.8	64.0	35	24	ABL51942	Ankyrin repeat mod
	29	12.8	64.0	36	21	AAA35872	Permutein linker e
	30	12.8	64.0	36	21	AAA35873	Permutein linker e
	31	12.8	64.0	51	22	AAL28412	Human SNP oligonuc
	32	12.8	64.0	53	17	AAT36510	Synthetic probe B
c	33	12.8	64.0	62	18	AAT47610	Primer pp65I for h
	34	12.8	64.0	70	13	AAQ32146	Probe specific to
c	35	12.8	64.0	70	13	AAQ32144	Probe specific to
c	36	12.6	63.0	20	18	AAT85719	Placental growth h
c	37	12.6	63.0	20	24	ABK46991	COX-2 antisense ol
	38	12.6	63.0	27	16	AAQ75152	GTI-II NEO cassett
	39	12.6	63.0	27	21	AAZ92601	Neomycin resistanc
c	40	12.6	63.0	30	22	AAS08319	Human CLASP-1, PCR
c	41	12.6	63.0	30	22	AAS07399	CLASP-1 PCR primer
c	42	12.6	63.0	30	22	AAH41906	Human CLASP-1 PCR
c	43	12.6	63.0	30	24	ABK84952	Cadherin-like asym
c	44	12.6	63.0	30	24	ABK84953	Cadherin-like asym
c	45	12.6	63.0	30	24	AAS18922	Human CLASP-1 anti

ALIGNMENTS

RESULT 1
AAZ31440
ID AAZ31440 standard; DNA; 20 BP.
XX
AC AAZ31440;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3610.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 4 A; 6 C; 8 G; 2 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GTGCCGACGTGGGACCCAGA 20
          ||||||||||||||||||||
Db      1 GTGCCGACGTGGGACCCAGA 20


RESULT 2
AAF97678/c
ID   AAF97678 standard; DNA; 21 BP.
XX
AC   AAF97678;
XX
DT   06-JUN-2001  (first entry)
XX
DE   Human gene single nucleotide polymorphism #2439.
XX
KW   Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
KW   polymorphism; vascular disease; coronary artery disease; forensics;
KW   myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
KW   pulmonary embolism; paternity test; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,A)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200118250-A2.
XX
PD   15-MAR-2001.
XX
PF   07-SEP-2000; 2000WO-US24503.
XX
PR   10-SEP-1999;   99US-0153357.
PR   26-JUL-2000; 2000US-0220947.
PR   16-AUG-2000; 2000US-0225724.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Lander ES,  Gargill M,  Ireland JS,  Bolk S,  Daley GQ,  McCarthy JJ;
XX
DR   WPI; 2001-226749/23.
XX
PT   Nucleic acids comprising single nucleotide polymorphisms, useful in
PT   applications such as forensics, paternity testing, medicine, genetic
PT   analysis and phenotype correlations to diseases such as diabetes and
PT   atherosclerosis -
XX
PS   Examples; Page 213; 242pp; English.
XX
CC   The present invention provides a method of diagnosing a vascular disease
CC   in an individual, involving determining the sequence at various
CC   polymorphic sites within the human thrombospondin 1 and thrombospondin 4
CC   genes. The sequences at a number of polymorphic sites are also provided
CC   in the specification. In particular, the method can be used in the
CC   diagnosis of atherosclerosis, myocardial infarction, coronary heart
CC   disease, stroke, peripheral vascular diseases, venous thromboembolism
CC   and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
CC   useful in forensics, paternity testing, genetic analysis and phenotype
CC   correlations to diseases. The present sequence is an example of one of
CC   the human gene SNPS shown in the specification.
XX
SQ   Sequence 21 BP; 3 A; 8 C; 6 G; 4 T; 0 other;

  Query Match             90.0%;  Score 18;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GTGCCGACGTGGGACCCA 18
          ||||||||||||||||||
Db     18 GTGCCGACGTGGGACCCA 1


RESULT 3
AAS07631
ID   AAS07631 standard; DNA; 27 BP.
XX
AC   AAS07631;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Streptococcus coelicolor mmyTOG promoter PCR primer CR344.
XX
KW   SCP1; ss; methylenomycin cluster; mmc; MmyR; MmfP; MmfH; mmyTOG;
KW   MmfL; MmfR; MmyT; MmyO; MmyG; MmYJ; Mmr; heterologous gene expression;
KW   PCR primer; CR344.
XX
OS   Streptococcus coelicolor.
XX
PN   WO200148228-A1.
XX
PD   05-JUL-2001.
XX
PF   20-DEC-2000; 2000WO-GB04972.
XX
PR   23-DEC-1999;   99GB-0030477.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Chater KF,  Bruton CJ,  O'Rouke SJ,  Wietzorrek AW;
XX
DR   WPI; 2001-425675/45.
XX
PT   Novel expression cassette for expressing a nucleic acid of interest,
PT   derived from the regulatory region of methylenomycin gene cluster of
PT   SCP1 plasmid of Streptomyces coelicolor A3(2)  -
XX
PS   Example 11; Page 75; 142pp; English.
XX
CC   The sequence represents a PCR primer used to amplify the promoter mmyTOG
CC   from an expression cassette SCP1 (the regulatory region of the
CC   methylenomycin cluster (mmc) from Streptomyces coelicolor A3(2), which
CC   encodes the MmyR, MmfP, MmfH, MmfL, MmfR, MmyT, MmyO, MmyG, MmYJ and
CC   partial Mmr polypeptides) for use as a strong promoter for heterologous
CC   gene expression. The expression cassette is useful for expressing a
CC   nucleic acid of interest, substantially only when the host cell culture
CC   reaches high cell density at or close to the stationary phase of host
CC   cell culture. In particular the system is useful in regulating
CC   methylenomycin production. Reduced or no expression of the nucleic acid
CC   of interest is observed earlier in growth, avoiding toxic effects of some
CC   gene products on growth and the system does not require addition of
CC   exogenous inducer. The methylenomycin cluster naturally present on a
```

```
CC   highly transmissible plasmid permits properly regulated expression in
CC   diverse Streptomyces host and the expression is driven by a strong
CC   promoter, leading to high yield of the desired end product.
XX
SQ   Sequence 27 BP; 3 A; 12 C; 9 G; 3 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 27;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GCCGACGTGGGACCCAG 19
           ||||||| ||||||| |
Db       3 GCCGACGCGGGACCCCG 19


RESULT 4
AAX83785/c
ID   AAX83785 standard; DNA; 28 BP.
XX
AC   AAX83785;
XX
DT   06-SEP-1999  (first entry)
XX
DE   PmeI restriction endonuclease PCR primer PmeI-IP1.
XX
KW   PmeI; restriction endonuclease; modification methylase;
KW   recombinant DNA technique; molecular cloning; gene characterization;
KW   Pseudomonas mendocina; PCR primer; ss.
XX
OS   Synthetic.
OS   Pseudomonas mendocina.
XX
PN   EP931835-A2.
XX
PD   28-JUL-1999.
XX
PF   24-NOV-1998;   98EP-0309629.
XX
PR   24-NOV-1997;   97US-0976703.
XX
PA   (NEWE ) NEW ENGLAND BIOLABS INC.
XX
PI   Chang Z,  Morgan RD;
XX
DR   WPI; 1999-397163/34.
XX
PT   New restriction endonuclease useful in recombinant DNA techniques
XX
PS   Example 1; Page 9; 28pp; English.
XX
CC   The present sequence represents a PCR primer for the 27 kD
CC   Pseudomonas mendocina PmeI restriction endonuclease protein. The PmeI
CC   restriction endonuclease is used to cut DNA in the middle of the site
CC   GTTTAAAC between the third T and the first A, to produce blunt ends.
CC   The PmeI restriction endonuclease is therefore useful in recombinant
CC   DNA techniques such as molecular cloning and gene characterization.
XX
SQ   Sequence 28 BP; 5 A; 7 C; 10 G; 6 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 20;  Length 28;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCA 18
           ||||||||| ||| |||
Db      19 TGCCGACGTCGGATCCA 3


RESULT 5
AAL34511/c
ID   AAL34511 standard; DNA; 50 BP.
XX
AC   AAL34511;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #7719.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 3620; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 7 A; 15 C; 17 G; 11 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 50;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GCCGACGTGGGACCCAG 19
           ||| | ||||||||||
Db      26 GCCCAAGTGGGACCCAG 10


RESULT 6
ABN29579
ID   ABN29579 standard; DNA; 65 BP.
XX
AC   ABN29579;
XX
DT   15-JUL-2002  (first entry)
XX
```

```
DE   Rat spliced transcript detection oligonucleotide SEQ ID NO:2327.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 2327; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 14 A; 14 C; 18 G; 19 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 65;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GCCGACGTGGGACCCAG 19
           || ||||||||| ||||
Db      23 GCAGACGTGGGAGCCAG 39


RESULT 7
AAX23277
ID   AAX23277 standard; DNA; 27 BP.
XX
AC   AAX23277;
XX
DT   11-JUN-1999  (first entry)
XX
DE   H. vulgare cysteine protease PCR primer GSP1.
XX
KW   Cysteine protease; barley; transgenic plant; disease resistance;
KW   prevention; treatment; infection; stress; temperature; drought; salt;
KW   wound; root-specific expression; antimicrobial; antistress;
KW   anti-wounding; PCR primer; ss.
XX
OS   Synthetic.
OS   Hordeum vulgare.
XX
PN   WO9910500-A1.
XX
PD   04-MAR-1999.
XX
PF   21-AUG-1998;   98WO-EP05339.
XX
PR   22-JAN-1998;   98DE-1002384.
PR   26-AUG-1997;   97DE-1037118.
XX
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI   Antoniw J,  Schlichter U,  Steinbiss H;
XX
DR   WPI; 1999-204668/17.
XX
PT   Nucleic acid encoding plant cysteine protease  - that increases
PT   resistance to injury and infection
XX
PS   Example 5; Page 71; 81pp; German.
XX
CC   This invention describes a novel Hordeum vulgare cysteine protease
CC   which is used to produce transgenic plants with (a) increased
CC   or reduced expression of protease or (b) increased resistance to disease
CC   (particularly prevention or treatment of infection) or other stress
CC   factors such as extreme temperatures, drought, salt or wounding. The
CC   invention also describes a regulatory region which is used to provide
CC   root-specific expression of heterologous DNA following infection/injury
CC   of roots. The products of the invention have antimicrobial, antistress
CC   and anti-wounding activity in plants.
XX
SQ   Sequence 27 BP; 6 A; 6 C; 11 G; 4 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 20;  Length 27;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCAGA 20
           |||  ||| || ||||||||
Db       6 GTGTGGACCTGCGACCCAGA 25


RESULT 8
AAI79205/c
ID   AAI79205 standard; DNA; 51 BP.
XX
AC   AAI79205;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:6146.
XX
KW   Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW   protein therapy; vaccine; probe; diagnostic assay; detection;
KW   quantitation; restorative therapy; polymorphic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200140521-A2.
XX
PD   07-JUN-2001.
XX
PF   30-NOV-2000; 2000WO-US32758.
XX
```

```
PR   30-NOV-1999;   99US-0168138.
PR   29-NOV-2000; 2000US-0726173.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-356160/37.
XX
PT   Polymorphic nucleic acid sequences, useful in genetic testing and
PT   therapy -
XX
PS   Claim 1; Page 2390; 2653pp; English.
XX
CC   AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC   sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC   AAM53114 to AAM53329 represent peptides related to human polymorphic
CC   polynucleotide sequences. The sequences can be used in gene and protein
CC   therapy, and in vaccine production. (I) and the polypeptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of polymorphic polypeptides.
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 51 BP; 14 A; 18 C; 12 G; 7 T; 0 other;

  Query Match            68.0%;  Score 13.6;  DB 22;  Length 51;
  Best Local Similarity  80.0%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCAGA 20
           || | |||||||| | ||||
Db      45 GTTCAGACGTGGGTCGCAGA 26


RESULT 9
ABN29322
ID   ABN29322 standard; DNA; 65 BP.
XX
AC   ABN29322;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Rat spliced transcript detection oligonucleotide SEQ ID NO:2070.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 2070; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 19 A; 16 C; 18 G; 12 T; 0 other;

  Query Match            68.0%;  Score 13.6;  DB 24;  Length 65;
  Best Local Similarity  80.0%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCAGA 20
           ||  |  ||||||||||||
Db       5 GTCACAGCGTGGGACCCAGA 24


RESULT 10
ABN29817/c
ID   ABN29817 standard; DNA; 65 BP.
XX
AC   ABN29817;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Rat spliced transcript detection oligonucleotide SEQ ID NO:2565.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
```

DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 2565; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 16 A; 23 C; 14 G; 12 T; 0 other;

Query Match 68.0%; Score 13.6; DB 24; Length 65;
Best Local Similarity 80.0%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGCCGACGTGGGACCCAGA 20
||||||| |||||| | ||
Db 20 GTGCCGATGTGGGAACTGGA 1

RESULT 11
AAL34194/c
ID AAL34194 standard; DNA; 50 BP.
XX
AC AAL34194;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #7402.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 3517; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 50 BP; 4 A; 26 C; 13 G; 7 T; 0 other;

Query Match 67.0%; Score 13.4; DB 22; Length 50;
Best Local Similarity 93.3%; Pred. No. 4e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTGCCGACGTGGGAC 15
||||||||||||| |
Db 29 GTGCCGACGTGGGGC 15

RESULT 12
AAX03341/c
ID AAX03341 standard; DNA; 27 BP.
XX
AC AAX03341;
XX
DT 26-MAR-1999 (first entry)
XX
DE PCR primer used to delete nucleotides 654-671 of HIV-1.
XX
KW Primer binding site; PBS; HIV-1; megaprimer; mutant; attenuated virus;
KW DNA polymerization; M184V mutated HIV-1 reverse transcriptase;
KW vaccine; acquired immunodeficiency syndrome; AIDS; PCR primer; ss.
XX
OS Synthetic.
OS Human immunodeficiency virus type 1.
XX
PN WO9900490-A2.
XX
PD 07-JAN-1999.
XX
PF 26-JUN-1998; 98WO-CA00617.
XX
PR 26-JUN-1997; 97CA-2208946.
XX
PA (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.
XX
PI Wainberg MA;
XX
DR WPI; 1999-095736/08.
XX

```
PT   New nucleic acid encoding an attenuated human immunodeficiency virus
PT   - containing a deletion downstream of the primer binding site and
PT   is useful as a vaccine against acquired immune deficiency syndrome
XX
PS   Example 1; Page 16; 41pp; English.
XX
CC   The present PCR primer is used (with primer AAX03301) to delete
CC   nucleotides 654-671 in sequences surrounding the primer binding site
CC   (PBS) of Human immunodeficiency virus type 1 (HIV-1). The amplification
CC   reactions produce nucleotide deleted mutant HIV-1 sequences. The
CC   specification describes deletion mutant viruses, where the deletion is
CC   located immediately downstream of PBS, and permits attenuation of the
CC   virus, as expressed by the PBS. More specifically,the deletion is within
CC   or includes either the 54 nucleotide segment immediately downstream of
CC   the PBS, or 16 nucleotide segment at the 3' end of the 54 nucleotide
CC   segment. The molecule also preferably includes the known M184V
CC   substitution. The deletion in the 54-nucleotide segment reduces the
CC   fidelity of the RNA- and DNA-dependent DNA polymerization of M184V
CC   mutated HIV-1 reverse transcriptase. The 54 nucleotide segment is needed
CC   for efficient reverse transcription of the viral DNA product, and
CC   efficient generation of viral mRNA, which affects subsequent viral
CC   protein synthesis and infectivity. The attenuated HIV (preferably HIV-1)
CC   forms a vaccine for immunizing a mammal (preferably a human) against
CC   HIV, which prevents acquired immunodeficiency syndrome (AIDS).
XX
SQ   Sequence 27 BP; 3 A; 7 C; 11 G; 6 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 20;  Length 27;
  Best Local Similarity   83.3%;  Pred. No. 5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCCGACGTGGGACCCAGA 20
           |||||   ||||||||||
Db      26 GCCGAACAGGGACCCAGA 9


RESULT 13
AAX89941
ID   AAX89941 standard; DNA; 33 BP.
XX
AC   AAX89941;
XX
DT   18-OCT-1999  (first entry)
XX
DE   Sequence of primer 1553-73.
XX
KW   Translational repression system; translational repressor; coat protein;
KW   constitutive promoter; heterologous gene; bacteriophage MS2; toxic gene;
KW   transcriptional control protein; motA; asiA; bacteriophage T4;
KW   PCR primer; protease gene; ss.
XX
OS   Synthetic.
OS   Bacteriophage sp.
XX
PN   WO9938985-A2.
XX
PD   05-AUG-1999.
XX
PF   27-JAN-1999;   99WO-US01725.
XX
PR   26-JAN-1999;   99US-0237712.
PR   28-JAN-1998;   98US-0072794.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Brown WC;
XX
DR   WPI; 1999-494095/41.
XX
PT   A bacteriophage MS2 translational repression system for use in
PT   cloning or expressing specific heterologous genes, especially toxic
PT   genes
XX
PS   Examples; Page 63; 133pp; English.
XX
CC   The invention relates to a translational repression system comprising a
CC   translational repressor operably linked to a constitutive promoter for
CC   use in cloning or expressing a specific heterologous gene. The
CC   translational repressor is bacteriophage MS2 coat protein. The improved
CC   expression vector systems are capable of expressing exogenous genes,
CC   including toxic genes, in Escherichia coli and other host cells. The
CC   bacteriophage MS2-based system can be used to stably clone and express
CC   certain toxic genes. The MS2-based T4 cassette is useful for the cloning
CC   of accessory proteins that are useful in the production of a target
CC   protein. The vector systems overcome the problem of promoter leakiness
CC   which leads to inappropriate transcription and expression of a gene
CC   cloned under the control of the promoter. This can be a problem when
CC   cloning toxic genes. The transcriptional control proteins, e.g. motA and
CC   asiA of bacteriophage T4, regulate transcription and provide a staged
CC   inducible promoter system which is much less complicated and more
CC   versatile than prior art non-staged systems. T4 middle promoters direct
CC   transcription from specific promoters while inhibiting transcription
CC   from E. coli promoters, which therefore minimises competition for
CC   translational apparatus and inhibits the cell from responding to target
CC   protein production by inducing transcription of protease genes.
CC   Sequences AAX89905-998 represent PCR primers used in the course of the
CC   invention.
XX
SQ   Sequence 33 BP; 4 A; 11 C; 11 G; 7 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 20;  Length 33;
  Best Local Similarity   83.3%;  Pred. No. 5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAG 19
           |||| ||||||| || ||
Db      12 TGCCTACGTGGGTCCTAG 29


RESULT 14
AAX89942/c
ID   AAX89942 standard; DNA; 33 BP.
XX
AC   AAX89942;
XX
DT   18-OCT-1999  (first entry)
XX
DE   Sequence of primer 1553-74.
XX
KW   Translational repression system; translational repressor; coat protein;
KW   constitutive promoter; heterologous gene; bacteriophage MS2; toxic gene;
KW   transcriptional control protein; motA; asiA; bacteriophage T4;
KW   PCR primer; protease gene; ss.
XX
OS   Synthetic.
OS   Bacteriophage sp.
XX
PN   WO9938985-A2.
XX
PD   05-AUG-1999.
XX
PF   27-JAN-1999;   99WO-US01725.
XX
PR   26-JAN-1999;   99US-0237712.
PR   28-JAN-1998;   98US-0072794.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Brown WC;
XX
DR   WPI; 1999-494095/41.
XX
PT   A bacteriophage MS2 translational repression system for use in
PT   cloning or expressing specific heterologous genes, especially toxic
PT   genes
XX
```

PS Examples; Page 63; 133pp; English.
XX
CC The invention relates to a translational repression system comprising a
CC translational repressor operably linked to a constitutive promoter for
CC use in cloning or expressing a specific heterologous gene. The
CC translational repressor is bacteriophage MS2 coat protein. The improved
CC expression vector systems are capable of expressing exogenous genes,
CC including toxic genes, in Escherichia coli and other host cells. The
CC bacteriophage MS2-based system can be used to stably clone and express
CC certain toxic genes. The MS2-based T4 cassette is useful for the cloning
CC of accessory proteins that are useful in the production of a target
CC protein. The vector systems overcome the problem of promoter leakiness
CC which leads to inappropriate transcription and expression of a gene
CC cloned under the control of the promoter. This can be a problem when
CC cloning toxic genes. The transcriptional control proteins, e.g. motA and
CC asiA of bacteriophage T4, regulate transcription and provide a staged
CC inducible promoter system which is much less complicated and more
CC versatile than prior art non-staged systems. T4 middle promoters direct
CC transcription from specific promoters while inhibiting transcription
CC from E. coli promoters, which therefore minimises competition for
CC translational apparatus and inhibits the cell from responding to target
CC protein production by inducing transcription of protease genes.
CC Sequences AAX89905-998 represent PCR primers used in the course of the
CC invention.
XX
SQ Sequence 33 BP; 8 A; 10 C; 10 G; 5 T; 0 other;

Query Match 66.0%; Score 13.2; DB 20; Length 33;
Best Local Similarity 83.3%; Pred. No. 5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TGCCGACGTGGGACCCAG 19
|||| ||||||| || ||
Db 26 TGCCTACGTGGGTCCTAG 9

RESULT 15
ABL32002/c
ID ABL32002 standard; DNA; 35 BP.
XX
AC ABL32002;
XX
DT 22-MAR-2002 (first entry)
XX
DE Dihydropyrimidine dehydrogenase (DPD) DNA fragment #1.
XX
KW Genetic analysis; sequence variant detection; genotype; ApoE; SNP;
KW single nucleotide polymorphism; polymorphic site; atherosclerosis;
KW dementia; Parkinson's disease; Huntington's disease; PCR primer;
KW neurodegenerative disease; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200190419-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US16577.
XX
PR 23-MAY-2000; 2000US-206613P.
PR 25-OCT-2000; 2000US-0696998.
PR 25-OCT-2000; 2000US-0697013.
PR 25-OCT-2000; 2000US-0697028.
XX
PA (VARI-) VARIAGENICS INC.
PA (STAN/) STANTON V P.
XX
PI Stanton VP;
XX
DR WPI; 2002-097670/13.
XX
PT Determining the haplotype of at least one allele of a selected gene at
PT two or more polymorphic sites, for assessing disease risk, comprises
PT allele-specific enrichment, optical mapping, or atomic force microscopy
PT -
XX
PS Example 1; Fig 22; 166pp; English.
XX
CC The present invention describes a method for determining the haplotype
CC of an allele of a selected gene at two or more polymorphic sites
CC comprising allele-specific enrichment, optical mapping, or atomic force
CC microscopy. The method is useful for genetic analysis when the DNA
CC segment being haplotyped consists of polymorphisms that are in some
CC degree of linkage disequilibrium with each other, that is they do not
CC assort randomly in the population being studied. The method allows early
CC implementation of preventive measures in patients at risk of diseases
CC such as atherosclerosis, dementia, Parkinson's disease, Huntington's
CC disease or other organic or vascular neurodegenerative diseases. Genotype
CC and haplotype information can be used to make diagnostic tests useful for
CC disease risk assessment, for prognostic prediction of the course or
CC outcome of a disease, to diagnose a disease or condition, or to select
CC an optimal therapy for a disease or condition. ABL31915 to ABL32035
CC represent nucleotide sequence used in the exemplification of the present
CC invention.
XX
SQ Sequence 35 BP; 9 A; 11 C; 6 G; 8 T; 1 other;

Query Match 66.0%; Score 13.2; DB 24; Length 35;
Best Local Similarity 83.3%; Pred. No. 5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GCCGACGTGGGACCCAGA 20
||||| |||| || ||||
Db 29 GCCGAAGTGGAACRCAGA 12

Search completed: November 23, 2002, 06:28:59
Job time : 102.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-10
Perfect score: 20
Sequence: 1 gtgccgacgtgggacccaga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	13.8	69.0	28	2	US-08-976-703-10	Sequence 10, Appl
	2	13.2	66.0	33	4	US-09-237-712-41	Sequence 41, Appl
c	3	13.2	66.0	33	4	US-09-237-712-42	Sequence 42, Appl
c	4	13.2	66.0	44	3	US-08-829-525-34	Sequence 34, Appl
c	5	13.2	66.0	44	4	US-08-609-583A-34	Sequence 34, Appl
c	6	13.2	66.0	44	4	US-08-937-399-34	Sequence 34, Appl
c	7	13.2	66.0	44	4	US-09-560-639-27	Sequence 27, Appl
c	8	13.2	66.0	44	4	US-09-310-367-34	Sequence 34, Appl
c	9	13.2	66.0	44	4	US-09-032-337-34	Sequence 34, Appl
	10	12.8	64.0	25	2	US-08-590-674-1	Sequence 1, Appli
c	11	12.8	64.0	62	2	US-08-658-665-153	Sequence 153, App
c	12	12.8	64.0	62	4	US-08-796-101-129	Sequence 129, App
c	13	12.8	64.0	62	4	US-09-085-273-153	Sequence 153, App
c	14	12.6	63.0	20	3	US-08-777-708C-14	Sequence 14, Appl
c	15	12.6	63.0	20	4	US-09-397-168-92	Sequence 92, Appl
	16	12.6	63.0	21	2	US-08-480-473B-61	Sequence 61, Appl
	17	12.6	63.0	21	3	US-08-915-213-61	Sequence 61, Appl
	18	12.6	63.0	21	4	US-09-235-217-61	Sequence 61, Appl
c	19	12.6	63.0	45	4	US-09-199-737-46	Sequence 46, Appl
c	20	12.6	63.0	45	4	US-09-058-333A-46	Sequence 46, Appl
c	21	12.2	61.0	20	4	US-09-561-497-65	Sequence 65, Appl
c	22	12.2	61.0	20	4	US-09-561-497-66	Sequence 66, Appl
	23	12.2	61.0	23	4	US-08-852-001-29	Sequence 29, Appl
	24	12.2	61.0	34	2	US-08-852-806-11	Sequence 11, Appl
	25	12.2	61.0	34	3	US-09-163-669-11	Sequence 11, Appl
	26	12.2	61.0	57	2	US-08-149-097D-25	Sequence 25, Appl
	27	12.2	61.0	92	3	US-08-973-965-4	Sequence 4, Appli
	28	12.2	61.0	93	3	US-08-973-965-5	Sequence 5, Appli
	29	12.2	61.0	93	3	US-08-973-965-38	Sequence 38, Appl
	30	12.2	61.0	97	3	US-08-973-965-40	Sequence 40, Appl
c	31	12	60.0	24	2	US-08-784-386-5	Sequence 5, Appli
c	32	12	60.0	25	2	US-08-784-386-10	Sequence 10, Appl
c	33	12	60.0	33	6	5520913-35	Patent No. 5520913
	34	12	60.0	44	1	US-08-344-695-5	Sequence 5, Appli
	35	12	60.0	50	4	US-09-298-886-19	Sequence 19, Appl
	36	12	60.0	50	4	US-09-298-886-25	Sequence 25, Appl
	37	12	60.0	50	4	US-09-298-886-28	Sequence 28, Appl
	38	12	60.0	54	1	US-08-311-486C-1026	Sequence 1026, Ap
	39	12	60.0	56	5	PCT-US95-02917-5	Sequence 5, Appli
	40	12	60.0	59	2	US-08-371-377-7	Sequence 7, Appli
c	41	12	60.0	67	4	US-09-025-769B-83	Sequence 83, Appl
	42	12	60.0	90	6	5175255-6	Patent No. 5175255
c	43	11.8	59.0	20	4	US-09-011-197-10	Sequence 10, Appl
c	44	11.8	59.0	25	2	US-08-933-749-2	Sequence 2, Appli
c	45	11.8	59.0	25	3	US-09-235-583-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-976-703-10/c
; Sequence 10, Application US/08976703
; Patent No. 5945288
; GENERAL INFORMATION:
; APPLICANT: CHANG, ZHIYUH
; APPLICANT: MORGAN, RICHARD D.
; TITLE OF INVENTION: METHOD FOR CLONING AND
; TITLE OF INVENTION: PRODUCING THE PmeI RESTRICTION ENDONUCLEASE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: New England Biolabs, Inc.
; STREET: 32 Tozer Road
; CITY: Beverly
; STATE: MA
; COUNTRY: US
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,703
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Gregory D
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-132
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 978-927-5054
; TELEFAX: 978-927-1705
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-976-703-10

Query Match 69.0%; Score 13.8; DB 2; Length 28;
Best Local Similarity 88.2%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       2 TGCCGACGTGGGACCCA 18
           ||||||||| ||| |||
Db      19 TGCCGACGTCGGATCCA 3


RESULT 2
US-09-237-712-41
; Sequence 41, Application US/09237712
; Patent No. 6180391
; GENERAL INFORMATION:
;  APPLICANT: BROWN, WILLIAM C.
;  TITLE OF INVENTION: HIGHLY EFFICIENT CONTROLLED EXPRESSION OF EXOGENOUS
;  TITLE OF INVENTION: GENES IN E. COLI
;  FILE REFERENCE: A-518
;  CURRENT APPLICATION NUMBER: US/09/237,712
;  CURRENT FILING DATE: 1999-01-26
;  EARLIER APPLICATION NUMBER: 60/072,794
;  EARLIER FILING DATE: 1998-01-28
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: oligonucleotide
US-09-237-712-41

  Query Match            66.0%;  Score 13.2;  DB 4;  Length 33;
  Best Local Similarity  83.3%;  Pred. No. 7.5e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAG 19
           |||| ||||||| || ||
Db      12 TGCCTACGTGGGTCCTAG 29


RESULT 3
US-09-237-712-42/c
; Sequence 42, Application US/09237712
; Patent No. 6180391
; GENERAL INFORMATION:
;  APPLICANT: BROWN, WILLIAM C.
;  TITLE OF INVENTION: HIGHLY EFFICIENT CONTROLLED EXPRESSION OF EXOGENOUS
;  TITLE OF INVENTION: GENES IN E. COLI
;  FILE REFERENCE: A-518
;  CURRENT APPLICATION NUMBER: US/09/237,712
;  CURRENT FILING DATE: 1999-01-26
;  EARLIER APPLICATION NUMBER: 60/072,794
;  EARLIER FILING DATE: 1998-01-28
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: oligonucleotide
US-09-237-712-42

  Query Match            66.0%;  Score 13.2;  DB 4;  Length 33;
  Best Local Similarity  83.3%;  Pred. No. 7.5e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAG 19
           |||| ||||||| || ||
Db      26 TGCCTACGTGGGTCCTAG 9


RESULT 4
US-08-829-525-34/c
; Sequence 34, Application US/08829525
; Patent No. 6084083
;  GENERAL INFORMATION:
;    APPLICANT:  Levinson, Douglas A.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036/2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/829,525
;      FILING DATE:  28-MAR-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/609,583
;      FILING DATE:  01-MAR-1996
;      APPLICATION NUMBER:  US 08/487,748
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/398,633
;      FILING DATE:  03-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7853-081
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-829-525-34

  Query Match            66.0%;  Score 13.2;  DB 3;  Length 44;
  Best Local Similarity  83.3%;  Pred. No. 7.4e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCA 18
           |||| ||| ||||| |||
Db      33 GTGCAGACTTGGGATCCA 16


RESULT 5
US-08-609-583A-34/c
; Sequence 34, Application US/08609583A
; Patent No. 6204371
;  GENERAL INFORMATION:
;    APPLICANT:  Levinson, Douglas A.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036/2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/609,583A
```

```
;      FILING DATE:  01-MAR-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/487,748
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/398,633
;      FILING DATE:  03-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7853-048
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-609-583A-34

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 44;
  Best Local Similarity   83.3%;  Pred. No. 7.4e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCA 18
           |||| ||| ||||| |||
Db      33 GTGCAGACTTGGGATCCA 16


RESULT 6
US-08-937-399-34/c
; Sequence 34, Application US/08937399
; Patent No. 6288218
;  GENERAL INFORMATION:
;    APPLICANT:  Levinson, Douglas A.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036/2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/937,399
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/609,583
;      FILING DATE:  01-MAR-1996
;      APPLICATION NUMBER:  US 08/487,748
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/398,633
;      FILING DATE:  03-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7853-048
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-937-399-34

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 44;
  Best Local Similarity   83.3%;  Pred. No. 7.4e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCA 18
           |||| ||| ||||| |||
Db      33 GTGCAGACTTGGGATCCA 16


RESULT 7
US-09-560-639-27/c
; Sequence 27, Application US/09560639
; Patent No. 6323334
; GENERAL INFORMATION:
;  APPLICANT: Kingsbury, G.
;  APPLICANT: Leiby, K.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION: TREATMENT OF IMMUNE DISORDERS
;  FILE REFERENCE: 7853-158
;  CURRENT APPLICATION NUMBER: US/09/560,639
;  CURRENT FILING DATE: 2000-04-28
;  EARLIER APPLICATION NUMBER: 60/155,862
;  EARLIER FILING DATE: 1999-09-24
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 27
;   LENGTH: 44
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 3' oligonucleotide
US-09-560-639-27

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 44;
  Best Local Similarity   83.3%;  Pred. No. 7.4e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCA 18
           |||| ||| ||||| |||
Db      33 GTGCAGACTTGGGATCCA 16


RESULT 8
US-09-310-367-34/c
; Sequence 34, Application US/09310367
; Patent No. 6414117
;  GENERAL INFORMATION:
;    APPLICANT:  Levinson, Douglas A.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036/2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/09/310,367
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/829,525
;      FILING DATE:  28-MAR-1997
;      APPLICATION NUMBER:  US 08/609,583
;      FILING DATE:  01-MAR-1996
;      APPLICATION NUMBER:  US 08/487,748
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/398,633
;      FILING DATE:  03-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7853-081
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-310-367-34

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 44;
  Best Local Similarity   83.3%;  Pred. No. 7.4e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCA 18
           |||| ||| ||||| |||
Db      33 GTGCAGACTTGGGATCCA 16


RESULT 9
US-09-032-337-34/c
; Sequence 34, Application US/09032337
; Patent No. 6455685
;  GENERAL INFORMATION:
;    APPLICANT:  Levinson, Douglas A.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036/2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/032,337
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/609,583
;      FILING DATE:  01-MAR-1996
;      APPLICATION NUMBER:  US 08/487,748
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/398,633
;      FILING DATE:  03-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7853-016
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-032-337-34

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 44;
  Best Local Similarity   83.3%;  Pred. No. 7.4e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCA 18
           |||| ||| ||||| |||
Db      33 GTGCAGACTTGGGATCCA 16


RESULT 10
US-08-590-674-1
; Sequence 1, Application US/08590674
; Patent No. 5849487
;  GENERAL INFORMATION:
;    APPLICANT:  Hase, Tetsu
;    APPLICANT:  Masubuchi, Harumi
;    TITLE OF INVENTION:  METHOD OF DETECTING A SPECIFIC
;    TITLE OF INVENTION:  POLYNUCLEOTIDE
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fitzpatrick, Cella, Harper & Scinto
;      STREET:  277 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  United States
;      ZIP:  10172-0194
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/590,674
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/356137
;      FILING DATE:  15-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wollman, Penina
;      REGISTRATION NUMBER:  30816
;      REFERENCE/DOCKET NUMBER:  1241.3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-2400
;      TELEFAX:  212-758-2982
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Synthetic DNA"
;    ORIGINAL SOURCE:
;      ORGANISM:  cytomegalovirus (CMV)
US-08-590-674-1

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+03;
```

```
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCGACGTGGGACCCAG 19
           ||||  ||||||||||
Db       3 CCGAAATGGGACCCAG 18


RESULT 11
US-08-658-665-153/c
; Sequence 153, Application US/08658665
; Patent No. 5997878
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Pincus, Steven E.
;    APPLICANT:  Cox, William I.
;    APPLICANT:  Kauffman, Elizabeth K.
;    TITLE OF INVENTION:  Recombinant Poxvirus - Cytomegalovirus,
;    TITLE OF INVENTION:  Compositions and Uses
;    NUMBER OF SEQUENCES:  190
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford, P.C.
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/658,665
;      FILING DATE:  05-JUN-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer Esq., William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2720.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)840-3333
;      TELEFAX:  (212)840-0712
;  INFORMATION FOR SEQ ID NO:  153:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  62 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-658-665-153

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 62;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCGACGTGGGACCCAG 19
           ||||  ||||||||||
Db      61 CCGAAATGGGACCCAG 46


RESULT 12
US-08-796-101-129/c
; Sequence 129, Application US/08796101
; Patent No. 6183752
;  GENERAL INFORMATION:
;    APPLICANT:  EPSTEIN, STEPHEN E.
;    APPLICANT:  FINKEL, TOREN
;    APPLICANT:  SPEIR, EDITH
;    APPLICANT:  ZHOU, YI FU
;    APPLICANT:  ZHU, JIANHUI
;    APPLICANT:  ERDILE, LORNE
;    APPLICANT:  PINCUS, STEVEN
;    TITLE OF INVENTION:  RESTENOSIS/ATHEROSCLEROSIS DIAGNOSIS,
;    TITLE OF INVENTION:  PROPHYLAXIS AND THERAPY
;    NUMBER OF SEQUENCES:  184
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CURTIS, MORRIS & SAFFORD, P.C.
;      STREET:  530 FIFTH AVENUE
;      CITY:  NEW YORK
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/796,101
;      FILING DATE:  05-FEB-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KOWALSKI, THOMAS J.
;      REGISTRATION NUMBER:  32,147
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 764-5574
;  INFORMATION FOR SEQ ID NO:  129:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  62 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-796-101-129

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 62;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCGACGTGGGACCCAG 19
           ||||  ||||||||||
Db      61 CCGAAATGGGACCCAG 46


RESULT 13
US-09-085-273-153/c
; Sequence 153, Application US/09085273
; Patent No. 6267965
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Pincus, Steven E.
;    APPLICANT:  Cox, William I.
;    APPLICANT:  Kauffman, Elizabeth K.
;    TITLE OF INVENTION:  RECOMBINANT POXVIRUS - CYTOMEGALOVIRUS,
;    TITLE OF INVENTION:  COMPOSITIONS AND USES
;    NUMBER OF SEQUENCES:  176
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/085,273
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/471,014
;      FILING DATE:  06-JUN-1995
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer Esq., William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2720
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  153:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  62 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-085-273-153

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 62;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCGACGTGGGACCCAG 19
           ||||  ||||||||||
Db      61 CCGAAATGGGACCCAG 46


RESULT 14
US-08-777-708C-14/c
; Sequence 14, Application US/08777708C
; Patent No. 6159934
; GENERAL INFORMATION:
;  APPLICANT:  Pescovitz, Ora H.
;  TITLE OF INVENTION:  USE OF GHRH-RP TO STIMULATE STEM CELL FACTOR PRODUCTION
;  FILE REFERENCE:  7037-175/IU-62
;  CURRENT APPLICATION NUMBER: US/08/777,708C
;  CURRENT FILING DATE:  1996-12-20
;  PRIOR APPLICATION NUMBER:  US 60/008,933
;  PRIOR FILING DATE:  1995-12-20
;  NUMBER OF SEQ ID NOS:  23
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-08-777-708C-14

  Query Match             63.0%;  Score 12.6;  DB 3;  Length 20;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCAG 19
           ||||  |  |||||| ||||
Db      19 GTGCAGTTGTGGGATCCAG 1


RESULT 15
US-09-397-168-92/c
; Sequence 92, Application US/09397168
; Patent No. 6344323
; GENERAL INFORMATION:
;  APPLICANT: Seifert, Wilfried
;  TITLE OF INVENTION: INHIBITION OF COX-2 EXPRESSION BY ANTISENSE OLIGONUCLEOTIDES
;  FILE REFERENCE: 11151/002001
;  CURRENT APPLICATION NUMBER: US/09/397,168
;  CURRENT FILING DATE: 1999-09-16
;  EARLIER APPLICATION NUMBER: US 60/100,590
;  EARLIER FILING DATE: 1998-09-16
;  NUMBER OF SEQ ID NOS: 92
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 92
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically generated antisense oligonucleotides
US-09-397-168-92

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 20;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAGA 20
           ||| ||| ||| || ||||
Db      20 TGCAGACCTGGTACTCAGA 2


Search completed: November 23, 2002, 06:36:11
Job time : 23.55 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-10
Perfect score: 20
Sequence: 1 gtgccgacgtgggacccaga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	13.2	66.0	35	10	US-09-967-013-60	Sequence 60, Appl
	2	13.2	66.0	35	10	US-09-967-013-61	Sequence 61, Appl
c	3	13.2	66.0	44	10	US-09-899-980A-27	Sequence 27, Appl
c	4	13.2	66.0	83	10	US-09-923-876-3414	Sequence 3414, Ap
c	5	12.6	63.0	30	10	US-09-736-969A-150	Sequence 150, App
c	6	12.6	63.0	30	10	US-09-736-960-130	Sequence 130, App
	7	12.4	62.0	50	10	US-09-880-732-15	Sequence 15, Appl
	8	12.4	62.0	51	10	US-09-880-732-14	Sequence 14, Appl
c	9	12.2	61.0	34	9	US-09-879-813-33	Sequence 33, Appl
	10	12.2	61.0	91	10	US-09-833-381-1832	Sequence 1832, Ap
	11	12.2	61.0	100	10	US-09-728-445-277	Sequence 277, App
c	12	12	60.0	25	10	US-09-866-108-3704	Sequence 3704, Ap
c	13	12	60.0	25	10	US-09-866-108-3705	Sequence 3705, Ap
c	14	12	60.0	25	10	US-09-866-108-3706	Sequence 3706, Ap
c	15	12	60.0	25	10	US-09-866-108-3707	Sequence 3707, Ap
c	16	12	60.0	25	10	US-09-866-108-3708	Sequence 3708, Ap
c	17	12	60.0	25	10	US-09-866-108-3709	Sequence 3709, Ap
	18	12	60.0	50	10	US-09-999-672-19	Sequence 19, Appl
	19	12	60.0	50	10	US-09-999-672-25	Sequence 25, Appl
	20	12	60.0	50	10	US-09-999-672-28	Sequence 28, Appl
	21	12	60.0	50	12	US-10-040-863-19	Sequence 19, Appl
	22	12	60.0	50	12	US-10-040-863-25	Sequence 25, Appl
	23	12	60.0	50	12	US-10-040-863-28	Sequence 28, Appl
	24	11.8	59.0	31	10	US-09-801-274-16	Sequence 16, Appl
c	25	11.8	59.0	96	10	US-09-864-761-32538	Sequence 32538, A
	26	11.6	58.0	22	10	US-09-999-672-1	Sequence 1, Appli
	27	11.6	58.0	22	12	US-10-040-863-1	Sequence 1, Appli
c	28	11.6	58.0	25	10	US-09-866-108-3710	Sequence 3710, Ap
c	29	11.6	58.0	25	10	US-09-866-108-3711	Sequence 3711, Ap
c	30	11.6	58.0	42	10	US-09-811-094-28	Sequence 28, Appl
c	31	11.6	58.0	42	10	US-09-810-644-28	Sequence 28, Appl
	32	11.6	58.0	50	10	US-09-999-672-16	Sequence 16, Appl
	33	11.6	58.0	50	10	US-09-999-672-22	Sequence 22, Appl
	34	11.6	58.0	50	12	US-10-040-863-16	Sequence 16, Appl
	35	11.6	58.0	50	12	US-10-040-863-22	Sequence 22, Appl
c	36	11.6	58.0	66	9	US-10-013-032-5	Sequence 5, Appli
c	37	11.6	58.0	78	10	US-09-972-809-6	Sequence 6, Appli
c	38	11.6	58.0	92	10	US-09-953-342-6	Sequence 6, Appli
	39	11.6	58.0	93	10	US-09-864-761-21595	Sequence 21595, A
	40	11.4	57.0	90	10	US-09-878-574-5301	Sequence 5301, Ap
	41	11.2	56.0	18	9	US-09-842-111-4	Sequence 4, Appli
	42	11.2	56.0	18	10	US-09-796-807-4	Sequence 4, Appli
c	43	11.2	56.0	27	10	US-09-371-307-29	Sequence 29, Appl
	44	11.2	56.0	67	10	US-09-847-637B-5	Sequence 5, Appli
	45	11	55.0	19	10	US-09-969-373-4348	Sequence 4348, Ap

ALIGNMENTS

RESULT 1
US-09-967-013-60/c
; Sequence 60, Application US/09967013
; Patent No. US20020045840A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr, Vincent P.
; TITLE OF INVENTION: METHOD FOR GENETIC ANALYSIS OF APOE DNA
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 11926-022001
; CURRENT APPLICATION NUMBER: US/09/967,013
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 60/206,613
; PRIOR FILING DATE: 2000-05-23
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 35
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: exemplary motif
US-09-967-013-60

Query Match 66.0%; Score 13.2; DB 10; Length 35;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       3 GCCGACGTGGGACCCAGA 20
           ||||| |||| || ||||
Db      29 GCCGAAGTGGAACRCAGA 12
```

RESULT 2
US-09-967-013-61
; Sequence 61, Application US/09967013
; Patent No. US20020045840A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr, Vincent P.
; TITLE OF INVENTION: METHOD FOR GENETIC ANALYSIS OF APOE DNA
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 11926-022001
; CURRENT APPLICATION NUMBER: US/09/967,013

```
;  CURRENT FILING DATE:  2000-10-25
;  PRIOR APPLICATION NUMBER: 60/206,613
;  PRIOR FILING DATE: 2000-05-23
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
;   LENGTH: 35
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: exemplary motif
US-09-967-013-61

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 35;
  Best Local Similarity   83.3%;  Pred. No. 5.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 GCCGACGTGGGACCCAGA 20
          ||||| |||| || ||||
Db      7 GCCGAAGTGGAACRCAGA 24


RESULT 3
US-09-899-980A-27/c
; Sequence 27, Application US/09899980A
; Patent No. US20020058800A1
; GENERAL INFORMATION:
;  APPLICANT: Kingsbury, G.
;  APPLICANT:  Leiby, K.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF IMMUNE DISORDERS
;  FILE REFERENCE: 7853-158
;  CURRENT APPLICATION NUMBER: US/09/899,980A
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/560,639
;  PRIOR FILING DATE: EARLIER FILING DATE: 2000-04-28
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 27
;   LENGTH: 44
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 3' oligonucleotide
US-09-899-980A-27

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 44;
  Best Local Similarity   83.3%;  Pred. No. 6e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCCGACGTGGGACCCA 18
          |||| ||| ||||| |||
Db     33 GTGCAGACTTGGGATCCA 16


RESULT 4
US-09-923-876-3414/c
; Sequence 3414, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
;  APPLICANT: Lalgudi, Raghunath V.
;  APPLICANT:  Kamigaki, Laura Y. (Ito)
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
;  FILE REFERENCE: PL-0012-1 CON
;  CURRENT APPLICATION NUMBER: US/09/923,876
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: 09/298,329
;  PRIOR FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: 60/085,331
;  PRIOR FILING DATE: 1998-05-05
;  NUMBER OF SEQ ID NOS: 6332
;  SOFTWARE: PERL Program
; SEQ ID NO 3414
;   LENGTH: 83
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020013958A1 700162221H1
;   NAME/KEY: unsure
;   LOCATION: 2, 28, 49, 63, 76, 79-80
;   OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-3414

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 83;
  Best Local Similarity   78.9%;  Pred. No. 6e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GTGCCGACGTGGGACCCAG 19
          | || ||||| ||| ||||
Db     50 GNGCGGACGTCGGAGCCAG 32


RESULT 5
US-09-736-969A-150/c
; Sequence 150, Application US/09736969A
; Patent No. US20020068302A1
; GENERAL INFORMATION:
;  APPLICANT: Lu, Peter
;  APPLICANT:  Garman, Jonathan David
;  APPLICANT:  Candia III, Albert Frederick
;  APPLICANT:  Arbor Vita Corporation
;  TITLE OF INVENTION: CLASP-4 Transmembrane Protein
;  FILE REFERENCE: 020054-000411US
;  CURRENT APPLICATION NUMBER: US/09/736,969A
;  CURRENT FILING DATE:  2000-12-13
;  PRIOR APPLICATION NUMBER: US 60/160,860
;  PRIOR FILING DATE: 1999-10-21
;  PRIOR APPLICATION NUMBER: US 60/162,498
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: US 60/170,453
;  PRIOR FILING DATE: 1999-12-13
;  PRIOR APPLICATION NUMBER: US 60/176,195
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: US 60/182,296
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: US 09/547,276
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,267
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,460
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,527
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 60/196,528
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: US 09/687,837
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US. 60/240,503
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,508
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,539
;  PRIOR FILING DATE: 2000-10-13
;  PRIOR APPLICATION NUMBER: US 60/240,543
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 153
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 150
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:antisense
;   OTHER INFORMATION: primer HClAS3'Kpn
```

US-09-736-969A-150

Query Match 63.0%; Score 12.6; DB 10; Length 30;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 GTGCCGACGTGGGACCCAG 19
           |||| || |||| ||| ||
Db      19 GTGCTGAAGTGGTACCAAG 1
```

RESULT 6
US-09-736-960-130/c
; Sequence 130, Application US/09736960
; Patent No. US20020102267A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter
; APPLICANT: Garman, Jonathan David
; APPLICANT: Candia III, Albert Frederick
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: CLASP-5 Transmembrane Protein
; FILE REFERENCE: 020054-000511US
; CURRENT APPLICATION NUMBER: US/09/736,960
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: US 60/160,860
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: US 60/162,498
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/170,453
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 60/176,195
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: US 60/182,296
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US 09/547,276
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,267
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,460
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,527
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/196,528
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 09/687,837
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,503
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,508
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,539
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/240,543
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 130
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:antisense
; OTHER INFORMATION: primer HClAS3'Kpn
US-09-736-960-130

Query Match 63.0%; Score 12.6; DB 10; Length 30;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 GTGCCGACGTGGGACCCAG 19
           |||| || |||| ||| ||
Db      19 GTGCTGAAGTGGTACCAAG 1
```

RESULT 7
US-09-880-732-15
; Sequence 15, Application US/09880732
; Patent No. US20020127561A1
; GENERAL INFORMATION:
; APPLICANT: GENICON SCIENCES CORPORATION
; APPLICANT: BEE, Gary
; APPLICANT: KOHNE, David E.
; APPLICANT: KORB, Linda
; APPLICANT: PETERSON, Todd
; APPLICANT: YGUERABIDE, Juan
; TITLE OF INVENTION: ASSAY FOR GENETIC POLYMORPHISMS USING SCATTERED LIGHT DETECTAB
; FILE REFERENCE: 089498/0403
; CURRENT APPLICATION NUMBER: US/09/880,732
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: US 60/210,988
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Exemplary probe for CYP2D6 allele detection
US-09-880-732-15

Query Match 62.0%; Score 12.4; DB 10; Length 50;
Best Local Similarity 92.9%; Pred. No. 1.5e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       6 GACGTGGGACCCAG 19
           ||| ||||||||||
Db      30 GACCTGGGACCCAG 43
```

RESULT 8
US-09-880-732-14
; Sequence 14, Application US/09880732
; Patent No. US20020127561A1
; GENERAL INFORMATION:
; APPLICANT: GENICON SCIENCES CORPORATION
; APPLICANT: BEE, Gary
; APPLICANT: KOHNE, David E.
; APPLICANT: KORB, Linda
; APPLICANT: PETERSON, Todd
; APPLICANT: YGUERABIDE, Juan
; TITLE OF INVENTION: ASSAY FOR GENETIC POLYMORPHISMS USING SCATTERED LIGHT DETECTAB
; FILE REFERENCE: 089498/0403
; CURRENT APPLICATION NUMBER: US/09/880,732
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: US 60/210,988
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Exemplary probe for CYP2D6 allele detection
US-09-880-732-14

Query Match 62.0%; Score 12.4; DB 10; Length 51;
Best Local Similarity 92.9%; Pred. No. 1.5e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       6 GACGTGGGACCCAG 19
           ||| ||||||||||
Db      31 GACCTGGGACCCAG 44
```

```
RESULT 9
US-09-879-813-33/c
; Sequence 33, Application US/09879813
; Patent No. US20020155453A1
; GENERAL INFORMATION:
;  APPLICANT: Sale, Julian E.
;  APPLICANT:  Neuberger, Michael S.
;  APPLICANT:  Cumbers, Sarah J.
;  TITLE OF INVENTION: Method of Generating Diversity
;  FILE REFERENCE: 18396/2005
;  CURRENT APPLICATION NUMBER: US/09/879,813
;  CURRENT FILING DATE:  2001-06-11
;  PRIOR APPLICATION NUMBER: 09/828,717
;  PRIOR FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: PCT/GB99/03358
;  PRIOR FILING DATE: 1999-10-08
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (16)..(29)
;   OTHER INFORMATION: U318
;   OTHER INFORMATION: The sequence  'ACGTCTGGGGCCA' is inserted and duplicates sequence
;   OTHER INFORMATION: s between positions 3 and 15
US-09-879-813-33

  Query Match             61.0%;  Score 12.2;  DB 9;  Length 34;
  Best Local Similarity   82.4%;  Pred. No. 1.8e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 CCGACGTGGGACCCAGA 20
           | ||||| || ||||||
Db      22 CAGACGTTGGCCCCAGA 6


RESULT 10
US-09-833-381-1832
; Sequence 1832, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1832
;   LENGTH: 91
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(91)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1832

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 91;
  Best Local Similarity   82.4%;  Pred. No. 1.9e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 CCGACGTGGGACCCAGA 20
           | ||| |||| ||||||
Db      74 CAGACCTGGGCCCCAGA 90


RESULT 11
US-09-728-445-277
; Sequence 277, Application US/09728445
; Patent No. US20020102543A1
; GENERAL INFORMATION:
;  APPLICANT: Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: No. US20020102543A1el Mutated Mammalian Cells and
;  TITLE OF INVENTION:  Animals
;  FILE REFERENCE: LEX-0102-USA
;  CURRENT APPLICATION NUMBER: US/09/728,445
;  CURRENT FILING DATE:  2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/168,358
;  PRIOR FILING DATE: 1999-12-01
;  NUMBER OF SEQ ID NOS: 891
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 277
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(100)
;   OTHER INFORMATION: n = A,T,C or G
US-09-728-445-277

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 100;
  Best Local Similarity   82.4%;  Pred. No. 1.9e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCA 18
             |||  ||||| ||||||
Db      65 TGCGTACGTGCGACCCA 81


RESULT 12
US-09-866-108-3704/c
; Sequence 3704, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 3704
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-3704

  Query Match             60.0%;  Score 12;  DB 10;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 2.3e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCAGA 20
           |||| | | ||||| ||| |
Db      25 GTGCTGGCCTGGGAGCCACA 6


RESULT 13
US-09-866-108-3705/c
; Sequence 3705, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 3705
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-3705

  Query Match             60.0%;  Score 12;  DB 10;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 2.3e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCAGA 20
           |||| | | ||||| ||| |
Db      24 GTGCTGGCCTGGGAGCCACA 5


RESULT 14
US-09-866-108-3706/c
; Sequence 3706, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 3706
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-3706

  Query Match             60.0%;  Score 12;  DB 10;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 2.3e+03;
```

Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 GTGCCGACGTGGGACCCAGA 20
           |||| | | ||||| ||| |
Db      23 GTGCTGGCCTGGGAGCCACA 4
```

RESULT 15
US-09-866-108-3707/c
; Sequence 3707, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 3707
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-3707

Query Match 60.0%; Score 12; DB 10; Length 25;
Best Local Similarity 75.0%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 GTGCCGACGTGGGACCCAGA 20
           |||| | | ||||| ||| |
Db      22 GTGCTGGCCTGGGAGCCACA 3
```

Search completed: November 23, 2002, 06:42:08
Job time : 18.25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-10
Perfect score: 20
Sequence: 1 gtgccgacgtgggacccaga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.2	71.0	74	17	BH215436	BH215436 1006027A1
	2	13.8	69.0	91	9	AA837667	AA837667 od73h03.s
	3	13.6	68.0	91	9	AA930982	AA930982 vy79f05.r
c	4	13.4	67.0	60	17	AZ776503	AZ776503 2M0010J11
c	5	13.2	66.0	50	9	AU103727	AU103727 AU103727
c	6	13.2	66.0	51	17	AZ438521	AZ438521 1M0228I12
	7	13.2	66.0	88	14	BQ756506	BQ756506 EBem09_SQ
c	8	13.2	66.0	95	17	AZ653293	AZ653293 1M0526G21
c	9	12.8	64.0	67	12	BG327801	BG327801 602426812
c	10	12.8	64.0	88	9	AU077180	AU077180 AU077180
	11	12.8	64.0	89	9	AA094815	AA094815 cp1650.se
c	12	12.6	63.0	50	9	AU106651	AU106651 AU106651
c	13	12.6	63.0	50	9	AU106652	AU106652 AU106652
c	14	12.6	63.0	50	9	AU106653	AU106653 AU106653
c	15	12.6	63.0	50	9	AU106654	AU106654 AU106654
c	16	12.6	63.0	50	9	AU106658	AU106658 AU106658
c	17	12.6	63.0	50	9	AU106659	AU106659 AU106659
c	18	12.6	63.0	50	9	AU106661	AU106661 AU106661
c	19	12.6	63.0	50	9	AU106662	AU106662 AU106662
c	20	12.6	63.0	50	9	AU106663	AU106663 AU106663
c	21	12.6	63.0	50	9	AU106664	AU106664 AU106664
c	22	12.6	63.0	50	9	AU106668	AU106668 AU106668
c	23	12.6	63.0	50	9	AU106670	AU106670 AU106670
c	24	12.6	63.0	50	9	AU106674	AU106674 AU106674
c	25	12.6	63.0	50	9	AU106675	AU106675 AU106675
c	26	12.6	63.0	50	9	AU106678	AU106678 AU106678
c	27	12.6	63.0	50	9	AU106679	AU106679 AU106679
c	28	12.6	63.0	50	9	AU106680	AU106680 AU106680
c	29	12.6	63.0	50	9	AU106681	AU106681 AU106681
c	30	12.6	63.0	50	9	AU106682	AU106682 AU106682
c	31	12.6	63.0	50	9	AU106683	AU106683 AU106683
c	32	12.6	63.0	50	9	AU106685	AU106685 AU106685
c	33	12.6	63.0	50	9	AU106686	AU106686 AU106686
c	34	12.6	63.0	50	9	AU106687	AU106687 AU106687
c	35	12.6	63.0	50	9	AU106688	AU106688 AU106688
c	36	12.6	63.0	50	9	AU106689	AU106689 AU106689
c	37	12.6	63.0	50	9	AU106690	AU106690 AU106690
c	38	12.6	63.0	50	9	AU106691	AU106691 AU106691
c	39	12.6	63.0	50	9	AU106692	AU106692 AU106692
c	40	12.6	63.0	50	9	AU106693	AU106693 AU106693
c	41	12.6	63.0	50	9	AU106694	AU106694 AU106694
c	42	12.6	63.0	50	9	AU106695	AU106695 AU106695
c	43	12.6	63.0	50	9	AU106696	AU106696 AU106696
c	44	12.6	63.0	50	9	AU106705	AU106705 AU106705
c	45	12.6	63.0	50	9	AU106708	AU106708 AU106708

ALIGNMENTS

RESULT 1
BH215436
LOCUS BH215436 74 bp DNA linear GSS 08-NOV-2001
DEFINITION 1006027A11.2EL_y2 1006 - RescueMu Grid G Zea mays genomic, DNA sequence.
ACCESSION BH215436
VERSION BH215436.1 GI:16806094
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 74)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1006027 row: 42
Class: transposon-tagged.
FEATURES Location/Qualifiers

```
     source          1. .74
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       15 a     24 c     25 g     10 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 74;
  Best Local Similarity   84.2%;  Pred. No. 1.3e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCCGACGTGGGACCCAG 19
          | |||||||||| ||| ||
Db     41 GCGCCGACGTGGCACCGAG 59


RESULT 2
AA837667
LOCUS       AA837667                 91 bp    mRNA    linear   EST 07-APR-1998
DEFINITION  od73h03.s1 NCI_CGAP_Ov2 Homo sapiens cDNA clone IMAGE:1373621
            similar to gb:M30938 LUPUS KU AUTOANTIGEN PROTEIN P86 (HUMAN);,
            mRNA sequence.
ACCESSION   AA837667
VERSION     AA837667.1  GI:2912866
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 258   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1373621"
                     /clone_lib="NCI_CGAP_Ov2"
                     /sex="female"
                     /tissue_type="ovary"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; mRNA made from invasive ovarian
                     tumor, cDNA made by oligo-dT priming.  Non-directionally
                     cloned. Size-selected on agarose gel, average insert  size
                     600 bp. Reference:  Krizman et al. (1996) Cancer Research
                     56:5380-5383."
BASE COUNT       21 a     27 c     29 g     14 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 91;
  Best Local Similarity   88.2%;  Pred. No. 2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 GCCGACGTGGGACCCAG 19
          | || ||||||||||||
Db     30 GACGGCGTGGGACCCAG 46


RESULT 3
AA930982
LOCUS       AA930982                 91 bp    mRNA    linear   EST 23-APR-1998
DEFINITION  vy79f05.r1 Stratagene mouse macrophage (#937306) Mus musculus cDNA
            clone IMAGE:1312449 5' similar to gb:X52839 60S RIBOSOMAL PROTEIN
            L17 (HUMAN);, mRNA sequence.
ACCESSION   AA930982
VERSION     AA930982.1  GI:3079415
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:682745
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 20.
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1312449"
                     /clone_lib="Stratagene mouse macrophage (#937306)"
                     /tissue_type="macrophage"
                     /dev_stage="WEHI-3 cell line"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: blood; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. WEHI-3 cell line. Average insert size: 1.5 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       25 a     19 c     25 g     22 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 91;
  Best Local Similarity   80.0%;  Pred. No. 2.5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GTGCCGACGTGGGACCCAGA 20
          |||| ||| || | ||||||
Db     21 GTGCAGACTTGTGGCCCAGA 40
```

```
RESULT 4
AZ776503/c
LOCUS       AZ776503                 60 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0010J11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0010J11 F, DNA sequence.
ACCESSION   AZ776503
VERSION     AZ776503.1  GI:12904143
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0010  row: J  column: 11
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 60.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0010J11"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       10 a     18 c     20 g     12 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 17;  Length 60;
  Best Local Similarity   93.3%;  Pred. No. 2.7e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACC 16
           |||| ||||||||||
Db      39 TGCCTACGTGGGACC 25
```

```
RESULT 5
AU103727/c
LOCUS       AU103727                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU103727 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP06017, mRNA sequence.
ACCESSION   AU103727
VERSION     AU103727.1  GI:13553248
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP06017"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT       12 a     14 c     16 g      8 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 50;
  Best Local Similarity   83.3%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCCGACGTGGGACCCAGA 20
           ||| .|||| |||| ||||
Db      30 GCCAACGTCGGACTCAGA 13
```

```
RESULT 6
AZ438521/c
LOCUS       AZ438521                 51 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0228I12R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0228I12 R, DNA sequence.
ACCESSION   AZ438521
VERSION     AZ438521.1  GI:10562534
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
```

```
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0228  row: I  column: 12
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 51.
FEATURES              Location/Qualifiers
     source           1. .51
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0228I12"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT       11 a     15 c     14 g     11 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAG 19
           |||| || |||| |||||
Db      31 TGCCCACATGGGCCCCAG 14


RESULT 7
BQ756506
LOCUS       BQ756506                 88 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  EBem09_SQ002_F10_R embryo, 1 Day germination, no treatment, cv
            Optic, EBem09 Hordeum vulgare cDNA clone EBem09_SQ002_F10 5', mRNA
            sequence.
ACCESSION   BQ756506
VERSION     BQ756506.1  GI:21964978
KEYWORDS    EST.
SOURCE      Hordeum vulgare.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
            Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
  TITLE     Development of Barley Transcriptome Resources
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Waugh R, Marshall DF
            Genome Dynamics/Computational Biology
            Scottish Crop Research Institute
            Invergowrie, Dundee, DD2 5DA, Scotland, UK
            Tel: 00 44 1382 562731
            Fax: 00 44 1382 562426
            Email: est@scri.sari.ac.uk.
FEATURES              Location/Qualifiers
     source           1. .88
                      /organism="Hordeum vulgare"
                      /cultivar="Optic"
                      /db_xref="taxon:4513"
                      /clone="EBem09_SQ002_F10"
                      /clone_lib="embryo, 1 Day germination, no treatment, cv
                      Optic, EBem09"
                      /tissue_type="embryo"
                      /dev_stage="1 Day germination"
                      /lab_host="DH10B"
                      /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
                      Non-normalised library, directionally cloned into pSPORT1.
                      Derived from embryos dissected from germinating grains (1
                      day) in glasshouse grown barley plants. Developed as part
                      of the barley transcriptome resources of BBSRC/SEERAD
                      funded cereal IGF (Investigating Gene Function) project."
BASE COUNT       20 a     28 c     25 g     15 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 14;  Length 88;
  Best Local Similarity   83.3%;  Pred. No. 3.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCA 18
           |||||||| || | ||||
Db      28 GTGCCGACCTGTGGCCCA 45


RESULT 8
AZ653293/c
LOCUS       AZ653293                 95 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0526G21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0526G21 R, DNA sequence.
ACCESSION   AZ653293
VERSION     AZ653293.1  GI:11790439
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0526  row: G  column: 21
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 95.
FEATURES              Location/Qualifiers
     source           1. .95
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0526G21"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
```

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 16 a 32 c 17 g 30 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 95;
Best Local Similarity 83.3%; Pred. No. 3.9e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       3 GCCGACGTGGGACCCAGA 20
           || ||| |||||| ||||
Db      74 GCAGACATGGGACACAGA 57
```

RESULT 9
BG327801/c
LOCUS BG327801 67 bp mRNA linear EST 27-FEB-2001
DEFINITION 602426812F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4564659 5', mRNA sequence.
ACCESSION BG327801
VERSION BG327801.1 GI:13134239
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 67)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM1280 row: h column: 04
High quality sequence stop: 67.
FEATURES Location/Qualifiers
source 1. .67
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4564659"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 21 a 21 c 13 g 12 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 12; Length 67;
Best Local Similarity 87.5%; Pred. No. 5.3e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       2 TGCCGACGTGGGACCC 17
           ||| ||| ||||||||
Db      64 TGCAGACCTGGGACCC 49
```

RESULT 10
AU077180/c
LOCUS AU077180 88 bp mRNA linear EST 04-MAY-2000
DEFINITION AU077180 Sugano cDNA library Homo sapiens cDNA clone Zrv61563 similar to 5'-end region of B.taurus mRNA for coatomer, mRNA sequence.
ACCESSION AU077180
VERSION AU077180.1 GI:7439794
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 88)
AUTHORS Suzuki,Y., Ishihara,D., Sasaki,M., Nakagawa,H., Hata,H., Tsunoda,T., Watanabe,M., Komatsu,T., Ota,T., Isogai,T., Suyama,A. and Sugano,S.
TITLE Statistical analysis of the 5' untranslated region of human mRNA using 'Oligo-Capped' cDNA libraries
JOURNAL Genomics 64 (3), 286-297 (2000)
MEDLINE 20221373
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997)
This clone was obtained from a '5'-end-enriched' cDNA library constructed by 'Oligo-Capping' method. The coding region starts from the 50 bp upstream to the 3'-end.
FEATURES Location/Qualifiers
source 1. .88
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Zrv61563"
/clone_lib="Sugano cDNA library"
/note="The cDNA was prepared using the anchor primer, H-T11G, from Genhunter"
BASE COUNT 16 a 26 c 24 g 22 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 9; Length 88;
Best Local Similarity 87.5%; Pred. No. 5.8e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 GTGCCGACGTGGGACC 16
           |||||||||||  ||
Db      23 GTGCCGACGTGGAGCC 8
```

RESULT 11
AA094815
LOCUS AA094815 89 bp mRNA linear EST 25-OCT-1996
DEFINITION cp1650.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA 5', mRNA sequence.
ACCESSION AA094815
VERSION AA094815.1 GI:1640408

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 89)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from fetal heart (1996)
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Liew CC
            Brigham and Women's Hospital
            Harvard Medical School
            75 Francis St. Boston, MA 02115, USA
            Tel: 6177328915
            Fax: 6179750995
            Email: cliew@rics.bwh.harvard.edu
            PCR PRimers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1. .89
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Human fetal heart, Lambda ZAP Express"
                     /lab_host="E. coli XL1-Blue"
                     /note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from human fetal hearts (8-10
                     weeks).  cDNA was synthesized using a XhoI-Oligo dT
                     adaptor-primer.  EcoRI adaptors were ligated, followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
BASE COUNT       19 a     27 c     22 g     21 t
ORIGIN

  Query Match           64.0%;  Score 12.8;  DB 9;  Length 89;
  Best Local Similarity   87.5%;  Pred. No. 5.8e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GCCGACGTGGGACCCA 18
           || ||||| |||||||
Db      40 GCTGACGTCGGACCCA 55


RESULT 12
AU106651/c
LOCUS       AU106651                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106651 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS03704, mRNA sequence.
ACCESSION   AU106651
VERSION     AU106651.1  GI:13556172
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CAS03704"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT       7 a     19 c     14 g     10 t
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 9;  Length 50;
  Best Local Similarity   78.9%;  Pred. No. 6e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAGA 20
           |||||||||| |  || ||
Db      31 TGCCGACGTGCGCTCCGGA 13


RESULT 13
AU106652/c
LOCUS       AU106652                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106652 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS03727, mRNA sequence.
ACCESSION   AU106652
VERSION     AU106652.1  GI:13556173
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CAS03727"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT       8 a     19 c     15 g      8 t
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 9;  Length 50;
  Best Local Similarity   78.9%;  Pred. No. 6e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAGA 20
           |||||||||| |  || ||
Db      28 TGCCGACGTGCGCTCCGGA 10


RESULT 14
AU106653/c
LOCUS       AU106653                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106653 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS04264, mRNA sequence.
ACCESSION   AU106653
```

```
VERSION     AU106653.1  GI:13556174
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CAS04264"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        6 a     19 c     14 g     11 t
ORIGIN

  Query Match            63.0%;  Score 12.6;  DB 9;  Length 50;
  Best Local Similarity  78.9%;  Pred. No. 6e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAGA 20
           ||||||||||  |  || ||
Db      31 TGCCGACGTGCGCTCCGGA 13


RESULT 15
AU106654/c
LOCUS       AU106654                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106654 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS04474, mRNA sequence.
ACCESSION   AU106654
VERSION     AU106654.1  GI:13556175
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CAS04474"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        7 a     19 c     14 g     10 t
ORIGIN

  Query Match            63.0%;  Score 12.6;  DB 9;  Length 50;
  Best Local Similarity  78.9%;  Pred. No. 6e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCCGACGTGGGACCCAGA 20
           ||||||||||  |  || ||
Db      31 TGCCGACGTGCGCTCCGGA 13


Search completed: November 26, 2002, 04:07:51
Job time : 758.8 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-11
Perfect score: 20
Sequence: 1 gacccccagggcactcatgg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	17.6	88.0	21	6	AX153915	AX153915 Sequence
c	2	15.2	76.0	44	6	I08796	I08796 Sequence 18
	3	13.6	68.0	59	6	AR073712	AR073712 Sequence
	4	13.6	68.0	61	6	AX085810	AX085810 Sequence
	5	13.6	68.0	93	6	E29108	E29108 Expression
	6	13.4	67.0	78	6	AX404028	AX404028 Sequence
c	7	13.2	66.0	24	6	AR104159	AR104159 Sequence
	8	13.2	66.0	51	6	AX117457	AX117457 Sequence
	9	13.2	66.0	51	6	AX158478	AX158478 Sequence
	10	13.2	66.0	51	9	AB013762	AB013762 Macaca as
	11	13.2	66.0	51	9	AB013763	AB013763 Macaca as
	12	13.2	66.0	51	9	AB013764	AB013764 Macaca fa
	13	13.2	66.0	67	9	S75425S04	S75432 collagen ty
	14	13.2	66.0	75	6	AR128284	AR128284 Sequence
	15	12.8	64.0	18	6	AR138054	AR138054 Sequence
c	16	12.8	64.0	60	9	HUMHPABA	M13192 Human hapto
c	17	12.8	64.0	66	6	E03396	E03396 DNA encodin
c	18	12.8	64.0	69	6	I66412	I66412 Sequence 10
	19	12.8	64.0	87	6	AX135351	AX135351 Sequence
c	20	12.8	64.0	99	10	AF357349	AF357349 Mus muscu
c	21	12.6	63.0	20	6	AX100702	AX100702 Sequence
c	22	12.6	63.0	21	6	AX096010	AX096010 Sequence
c	23	12.6	63.0	24	6	AX291553	AX291553 Sequence
c	24	12.6	63.0	28	6	AR182413	AR182413 Sequence
c	25	12.6	63.0	28	6	AR182415	AR182415 Sequence
c	26	12.6	63.0	29	6	AR199662	AR199662 Sequence
	27	12.6	63.0	30	6	AR068885	AR068885 Sequence
c	28	12.6	63.0	40	6	AR033896	AR033896 Sequence
c	29	12.6	63.0	40	6	AR175029	AR175029 Sequence
c	30	12.6	63.0	40	6	AX032461	AX032461 Sequence
	31	12.6	63.0	45	6	AR033949	AR033949 Sequence
	32	12.6	63.0	45	6	AR175082	AR175082 Sequence
	33	12.6	63.0	45	6	AX032526	AX032526 Sequence
c	34	12.6	63.0	51	6	AX162480	AX162480 Sequence
c	35	12.6	63.0	51	6	AX190058	AX190058 Sequence
	36	12.6	63.0	74	9	HUM13COL22	M68991 Human alpha
c	37	12.6	63.0	76	9	AF040224	AF040224 Homo sapi
	38	12.6	63.0	76	9	S75385S18	S75420 collagen ty
	39	12.4	62.0	18	6	I25834	I25834 Sequence 14
c	40	12.4	62.0	21	6	I05985	I05985 Sequence 2
c	41	12.4	62.0	21	6	I08790	I08790 Sequence 6
	42	12.4	62.0	24	6	AR091297	AR091297 Sequence
	43	12.4	62.0	24	6	AR137679	AR137679 Sequence
c	44	12.4	62.0	37	6	AR171242	AR171242 Sequence
	45	12.4	62.0	38	6	AR058104	AR058104 Sequence

ALIGNMENTS

RESULT 1
AX153915/c
LOCUS AX153915 21 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 13 from Patent WO0138576.
ACCESSION AX153915
VERSION AX153915.1 GI:14535529
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 21)
AUTHORS Cargill,M., Ireland,J.S. and Lander,E.S.
TITLE Human single nucleotide polymorphisms
JOURNAL Patent: WO 0138576-A 13 31-MAY-2001;

```
          WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        3 a      5 c      9 g      3 t      1 others
ORIGIN

  Query Match             88.0%;  Score 17.6;  DB 6;  Length 21;
  Best Local Similarity   94.4%;  Pred. No. 6.6e+02;
  Matches   17;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 CCCCCAGGGCACTCATGG 20
           ||||||||||:|||||||
Db      21 CCCCCAGGGCRCTCATGG 4


RESULT 2
I08796/c
LOCUS       I08796                    44 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 18 from Patent WO 8804690.
ACCESSION   I08796
VERSION     I08796.1  GI:588501
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Bollen,A.J., Gheysen,D., Jacobs,P., Pierard,L. and Collen,D.J.
  JOURNAL   Patent: WO 8804690-A 18 30-JUN-1988;
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="unknown"
BASE COUNT       10 a     10 c     14 g     10 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 6;  Length 44;
  Best Local Similarity   85.0%;  Pred. No. 9.1e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           ||||| ||| ||| ||||||
Db      42 GACCCTCAGAGCAGTCATGG 23


RESULT 3
AR073712
LOCUS       AR073712                  59 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 5 from patent US 5952190.
ACCESSION   AR073712
VERSION     AR073712.1  GI:10000472
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Joenje,H. and Lo Ten Foe,J.R.
  TITLE     cDNA for fanconi anemia complementation group A
  JOURNAL   Patent: US 5952190-A 5 14-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="unknown"
BASE COUNT       18 a     16 c     12 g     13 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 59;
  Best Local Similarity   80.0%;  Pred. No. 5.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           |||| ||| |||  ||||||
Db      27 GACCACCATGGCCATCATGG 46


RESULT 4
AX085810
LOCUS       AX085810                  61 bp    DNA     linear   PAT 09-MAR-2001
DEFINITION  Sequence 72 from Patent WO0112858.
ACCESSION   AX085810
VERSION     AX085810.1  GI:13275760
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   He,T.C., Kinzler,K.W. and Vogelstein,B.
  TITLE     ppar_g(d) links apc to chemopreventive drugs
  JOURNAL   Patent: WO 0112858-A 72 22-FEB-2001;
            The Johns Hopkins University (US)
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       20 a     16 c     14 g     11 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 61;
  Best Local Similarity   80.0%;  Pred. No. 5.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           |||| ||| ||||  |||||
Db      28 GACCACCATGGCAACCATGG 47


RESULT 5
E29108
LOCUS       E29108                    93 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Expression regulatory region of stress protein HSP47 and
            utilization thereof.
ACCESSION   E29108
VERSION     E29108.1  GI:13020964
KEYWORDS    JP 1999243968-A/1.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Kazuhiro,N.
  TITLE     Expression regulatory region of stress protein HSP47 and
            utilization thereof
  JOURNAL   Patent: JP 1999243968-A 1 14-SEP-1999;
            SCIENCE & TECH AGENCY
COMMENT     OS   Unidentified
            PN   JP 1999243968-A/1
            PD   14-SEP-1999
            PF   05-MAR-1998 JP 1998071489
            PR
            PI   KAZUHIRO NAGATA
            PC   C12N15/09,A61K31/70,A61K48/00//(C12N15/09,C12R1:91),C12N15/00,
            PC   (C12N15/00,C12R1:91)
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .93
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       18 a     26 c     29 g     20 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 93;
  Best Local Similarity   80.0%;  Pred. No. 5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy      1 GACCCCCAGGGCACTCATGG 20
          | |  ||||||||| |||||
Db      6 GGCATCCAGGGCACGCATGG 25


RESULT 6
AX404028
LOCUS       AX404028                  78 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 25 from Patent EP1195161.
ACCESSION   AX404028
VERSION     AX404028.1  GI:21437354
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1
  AUTHORS   Morsey,M.A., Sheppard,M.G. and Wheeler,D.W.
  TITLE     Anti-ige vaccines
  JOURNAL   Patent: EP 1195161-A 25 10-APR-2002;
            Pfizer Products Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
BASE COUNT       12 a     36 c     22 g      8 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 78;
  Best Local Similarity   93.3%;  Pred. No. 6.5e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 CCCAGGGCACTCATG 19
          ||||||||| ||||||
Db     25 CCCAGGGCCCTCATG 39


RESULT 7
AR104159/c
LOCUS       AR104159                  24 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 8 from patent US 6093542.
ACCESSION   AR104159
VERSION     AR104159.1  GI:12816867
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Romano,J., Pal,R. and Shurtliff,R.
  TITLE     Isothermal transcription based amplification assay for the
            detection and quantitation of macrophage derived chemokine RNA
  JOURNAL   Patent: US 6093542-A 8 25-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        4 a      8 c      8 g      4 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 9.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 CCCCCAGGGCACTCATGG 20
          | ||||||||||||  ||
Db     23 CACCCAGGGCACTCTGGG 6


RESULT 8
AX117457
LOCUS       AX117457                  51 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 2580 from Patent WO0129262.
ACCESSION   AX117457
VERSION     AX117457.1  GI:14034408
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Picoult-Newburg,L. and Pohl,M.
  TITLE     Genotyping reagents, kits and methods of use thereof
  JOURNAL   Patent: WO 0129262-A 2580 26-APR-2001;
            Orchid BioSciences, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       18 a     14 c     10 g      9 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 8.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ACCCCCAGGGCACTCATG 19
          |||||||||  || ||||
Db     14 ACCCCCAGGCGACACATG 31


RESULT 9
AX158478
LOCUS       AX158478                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 1806 from Patent WO0140521.
ACCESSION   AX158478
VERSION     AX158478.1  GI:14539809
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 1806 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (1805 is other entry)
                     Accession number cg35066497"
BASE COUNT        7 a     18 c     19 g      7 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 8.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GACCCCCAGGGCACTCAT 18
          || |||||||| |||| |
Db     25 GAGCCCCAGGGTACTCCT 42


RESULT 10
AB013762
LOCUS       AB013762                  51 bp    DNA     linear   PRI 29-JUN-2002
DEFINITION  Macaca assamensis gene for immunoglobulin alpha heavy chain,
            partial cds, isolate:UTS400.
ACCESSION   AB013762
VERSION     AB013762.1  GI:6681477
KEYWORDS    IgA.
SOURCE      Macaca assamensis (isolate:UTS400) DNA.
  ORGANISM  Macaca assamensis
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Sumiyama,K., Saitou,N. and Ueda,S.
  TITLE     Adaptive evolution of the IgA hinge region in primates
  JOURNAL   Mol. Biol. Evol. 19 (7), 1093-1099 (2002)
  MEDLINE   22077633
   PUBMED   12082128
REFERENCE   2  (bases 1 to 51)
  AUTHORS   Sumiyama,K., Ueda,S. and Saitou,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-1998) Kenta Sumiyama, National Institute of
            Genetics, Evolutionary Genetics Division; Yata 1111, Mishima,
            Shizuoka 411-8540, Japan (E-mail:ksumiyam@lab.nig.ac.jp,
            Tel:+81-559-81-6790, Fax:+81-559-81-6794)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Macaca assamensis"
                     /isolate="UTS400"
                     /db_xref="taxon:9551"
     gene            1. .51
                     /gene="C alpha"
     CDS             <1. .>51
                     /gene="C alpha"
                     /codon_start=3
                     /product="immunoglobulin alpha heavy chain"
                     /protein_id="BAA88725.1"
                     /db_xref="GI:6681478"
                     /db_xref="IMGT/LIGM:AB013762"
                     /translation="LQFPQSHPHVPHAASP"
BASE COUNT        9 a     25 c      8 g      9 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 8.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ACCCCCAGGGCACTCATG 19
          ||||||| | | ||||||
Db     22 ACCCCCATGTCCCTCATG 39


RESULT 11
AB013763
LOCUS       AB013763                  51 bp    DNA     linear   PRI 29-JUN-2002
DEFINITION  Macaca assamensis gene for immunoglobulin alpha heavy chain,
            partial cds, isolate:UTS398B.
ACCESSION   AB013763
VERSION     AB013763.1  GI:6681479
KEYWORDS    IgA.
SOURCE      Macaca assamensis (isolate:UTS398B) DNA.
  ORGANISM  Macaca assamensis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Sumiyama,K., Saitou,N. and Ueda,S.
  TITLE     Adaptive evolution of the IgA hinge region in primates
  JOURNAL   Mol. Biol. Evol. 19 (7), 1093-1099 (2002)
  MEDLINE   22077633
   PUBMED   12082128
REFERENCE   2  (bases 1 to 51)
  AUTHORS   Sumiyama,K., Ueda,S. and Saitou,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-1998) Kenta Sumiyama, National Institute of
            Genetics, Evolutionary Genetics Division; Yata 1111, Mishima,
            Shizuoka 411-8540, Japan (E-mail:ksumiyam@lab.nig.ac.jp,
            Tel:+81-559-81-6790, Fax:+81-559-81-6794)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Macaca assamensis"
                     /isolate="UTS398B"
                     /db_xref="taxon:9551"
     gene            1. .51
                     /gene="C alpha"
     CDS             <1. .>51
                     /gene="C alpha"
                     /codon_start=3
                     /product="immunoglobulin alpha heavy chain"
                     /protein_id="BAA88726.1"
                     /db_xref="GI:6681480"
                     /db_xref="IMGT/LIGM:AB013763"
                     /translation="LQFPQSHPHVPHAASP"
BASE COUNT        9 a     25 c      8 g      9 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 8.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ACCCCCAGGGCACTCATG 19
          ||||||| | | ||||||
Db     22 ACCCCCATGTCCCTCATG 39


RESULT 12
AB013764
LOCUS       AB013764                  51 bp    DNA     linear   PRI 29-JUN-2002
DEFINITION  Macaca fascicularis gene for immunoglobulin alpha heavy chain,
            partial cds, isolate:UTS362B.
ACCESSION   AB013764
VERSION     AB013764.1  GI:6681481
KEYWORDS    IgA.
SOURCE      Macaca fascicularis (isolate:UTS362B) DNA.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   Sumiyama,K., Saitou,N. and Ueda,S.
  TITLE     Adaptive evolution of the IgA hinge region in primates
  JOURNAL   Mol. Biol. Evol. 19 (7), 1093-1099 (2002)
  MEDLINE   22077633
   PUBMED   12082128
REFERENCE   2  (bases 1 to 51)
  AUTHORS   Sumiyama,K., Ueda,S. and Saitou,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-1998) Kenta Sumiyama, National Institute of
            Genetics, Evolutionary Genetics Division; Yata 1111, Mishima,
            Shizuoka 411-8540, Japan (E-mail:ksumiyam@lab.nig.ac.jp,
            Tel:+81-559-81-6790, Fax:+81-559-81-6794)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Macaca fascicularis"
                     /isolate="UTS362B"
                     /db_xref="taxon:9541"
     gene            1. .51
                     /gene="C alpha"
     CDS             <1. .>51
                     /gene="C alpha"
                     /codon_start=3
                     /product="immunoglobulin alpha heavy chain"
                     /protein_id="BAA88727.1"
                     /db_xref="GI:6681482"
                     /db_xref="IMGT/LIGM:AB013764"
                     /translation="LQFPQSHPHVPHAASP"
BASE COUNT        9 a     25 c      8 g      9 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 8.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ACCCCCAGGGCACTCATG 19
          ||||||| | | ||||||
```

```
Db     22 ACCCCCATGTCCCTCATG 39


RESULT 13
S75425S04
LOCUS       S75425S04                67 bp    DNA     linear   PRI 02-JUN-2000
DEFINITION  collagen type VI alpha 2(VI) chain (exons 1-19 and intron/exon
            junctions) [human, Genomic, 67 nt, segment 4 of 19].
ACCESSION   S75432
VERSION     S75432.1  GI:241989
KEYWORDS    .
SEGMENT     4 of 19
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Saitta,B., Wang,Y.M., Renkart,L., Zhang,R.Z., Pan,T.C., Timpl,R.
            and Chu,M.L.
  TITLE     The exon organization of the triple-helical coding regions of the
            human alpha 1(VI) and alpha 2(VI) collagen genes is highly similar
  JOURNAL   Genomics 11 (1), 145-153 (1991)
  MEDLINE   92112205
   PUBMED   1765372
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 75432] from the original journal article.
            This sequence comes from Figure 4.
            Map location: chromosome 21.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     exon            31. .57
                     /number=4
BASE COUNT       12 a     23 c     17 g     15 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 9;  Length 67;
  Best Local Similarity  83.3%;  Pred. No. 8.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 CCCCCAGGGCACTCATGG 20
          ||||||||||  || |||
Db     24 CCCCCAGGGCGTTCCTGG 41


RESULT 14
AR128284
LOCUS       AR128284                 75 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 146 from patent US 6183957.
ACCESSION   AR128284
VERSION     AR128284.1  GI:14115946
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Cole,S., Buchrieser-Brosch,R., Gordon,S. and Billault,A.
  TITLE     Method for isolating a polynucleotide of interest from the genome
            of a mycobacterium using a BAC-based DNA library application to the
            detection of mycobacteria
  JOURNAL   Patent: US 6183957-A 146 06-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="unknown"
BASE COUNT       13 a     25 c     23 g     14 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 6;  Length 75;
  Best Local Similarity  83.3%;  Pred. No. 8.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ACCCCCAGGGCACTCATG 19
          |||||   ||||||||||
Db     30 ACCCCACAGGCACTCATG 47


RESULT 15
AR138054
LOCUS       AR138054                 18 bp    DNA     linear   PAT 16-JUN-2001
DEFINITION  Sequence 64 from patent US 6197584.
ACCESSION   AR138054
VERSION     AR138054.1  GI:14479563
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 18)
  AUTHORS   Bennett,C.Frank. and Cowsert,L.M.
  TITLE     Antisense modulation of CD40 expression
  JOURNAL   Patent: US 6197584-A 64 06-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="unknown"
BASE COUNT        4 a      8 c      4 g      2 t
ORIGIN

  Query Match            64.0%;  Score 12.8;  DB 6;  Length 18;
  Best Local Similarity  87.5%;  Pred. No. 1.6e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 ACCCCCAGGGCACTCA 17
          ||| ||||||| ||||
Db      2 ACCACCAGGGCTCTCA 17


Search completed: December  3, 2002, 18:13:11
Job time : 310.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-11
Perfect score: 20
Sequence: 1 gacccccagggcactcatgg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	21	AAZ31441	Human neuropilin m
2	c	16.4	82.0	21	22	AAH62112	Neuropilin 1 (NRP1
3		16	80.0	20	21	AAZ31450	Human neuropilin m
4		15.2	76.0	51	22	AAL27306	Human SNP oligonuc
5	c	15.2	76.0	60	24	ABN34554	Human spliced tran
6	c	14.4	72.0	39	24	ABK88692	Human transmembran
7	c	14.4	72.0	39	24	ABK13410	Tumour necrosis fa
8		14.2	71.0	51	22	AAL27305	Human SNP oligonuc
9		14.2	71.0	51	22	AAL27307	Human SNP oligonuc
10	c	14.2	71.0	60	24	ABN59176	Human spliced tran
11		13.8	69.0	51	22	AAL27308	Human SNP oligonuc
12	c	13.6	68.0	51	22	AAL28078	Human SNP oligonuc
13		13.6	68.0	59	19	AAV18190	Primer for Fanconi
14	c	13.6	68.0	60	24	ABN39793	Human spliced tran
15		13.6	68.0	61	22	AAF75272	Human PPARgamma PC
16		13.6	68.0	93	20	AAZ27697	Intron from HSP47
17	c	13.2	66.0	24	20	AAX90284	Macrophage derived
18	c	13.2	66.0	31	20	AAX07235	PCR primer ZC15,33
19	c	13.2	66.0	38	21	AAC82076	p53 PCR primer FP5
20	c	13.2	66.0	50	22	AAL28077	Human SNP oligonuc
21		13.2	66.0	51	22	AAI74865	Human silent SNP c
22		13.2	66.0	51	22	AAH39784	Human SNP flanking
23		13.2	66.0	60	24	ABN46919	Human spliced tran
24		13.2	66.0	75	21	ABQ62647	Mycobacterium tube
25	c	13.2	66.0	80	22	AAC83397	Primer luci2. Syn
26		13.2	66.0	82	22	AAC83396	Primer luci1. Syn
27	c	13	65.0	18	14	AAQ40337	p58 nested primer
28		13	65.0	20	24	ABQ61256	Human aquaporin 5
29	c	13	65.0	21	19	AAZ26653	Human polymorphic
30	c	13	65.0	21	22	AAF96423	Human gene single
31		13	65.0	50	22	AAL30652	Human SNP oligonuc
32		12.8	64.0	18	20	AAZ40915	Human CD40 phospho
33		12.8	64.0	18	21	AAZ47748	Human CD40 antisen
34		12.8	64.0	20	21	AAC86393	Primer BMV R. Uni
35		12.8	64.0	20	22	AAC91199	BMV-F primer. Syn
36		12.8	64.0	20	22	AAF23596	Primer BMV R. Uni
37		12.8	64.0	20	24	AAS20531	Purified viral RNA
38		12.8	64.0	25	21	AAZ45010	UDP-glucuronosyltr
39		12.8	64.0	25	22	AAS43234	Human Oestrogen re
40	c	12.8	64.0	27	19	AAV19570	Human LDL receptor
41	c	12.8	64.0	31	16	AAQ85848	MHC class I allele
42	c	12.8	64.0	31	16	AAQ85778	MHC CW3 amplificat
43	c	12.8	64.0	31	16	AAQ85742	MHC CW3 cDNA nts 1
44	c	12.8	64.0	33	24	ABL42117	Primer #3 for huma
45		12.8	64.0	47	21	AAZ65851	Human map-related

ALIGNMENTS

RESULT 1
AAZ31441
ID AAZ31441 standard; DNA; 20 BP.
XX
AC AAZ31441;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3611.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 4 A; 8 C; 6 G; 2 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 4.6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           ||||||||||||||||||||
Db       1 GACCCCCAGGGCACTCATGG 20


RESULT 2
AAH62112/c
ID   AAH62112 standard; DNA; 21 BP.
XX
AC   AAH62112;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Neuropilin 1 (NRP1) polymorphism containing DNA fragment #13.
XX
KW   Single nucleotide polymorphism; SNP; human; cancer; inflammation;
KW   heart disease; paternity testing; forensic science; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,T)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200138576-A2.
XX
PD   31-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US31639.
XX
PR   24-NOV-1999;   99US-0167334.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-367705/38.
XX
PT   New nucleic acid segments of the human genome, particularly from genes
PT   including polymorphic sites,for phenotype correlation, forensics,
PT   paternity testing, medicine and genetic analysis  -
XX
PS   Claim 1; Page 29; 80pp; English.
XX
CC   DNA sequences AAH62100 - AAH62688 represent segments of human genes which
CC   contain single nucleotide polymorphisms (SNPs). A method is included in
CC   the invention for analysing a nucleic acid sample, which consists of
CC   determining the base occupying any one of the polymorphic sites given in
CC   the SNP containing sequences. The nucleotide sequences can be used in the
CC   diagnosis or monitoring of diseases, such as cancer, inflammation, heart
CC   diseases, diseases of the cardiovascular system, and infection by
CC   microorganisms. The oligonucleotides are also useful in the manufacture
CC   of a medicament for the treatment or prophylaxis of the diseases, and as
CC   a pharmaceutical. SNP containing oligonucleotides are useful in
CC   applications such as phenotype correlation, forensics, paternity testing,
CC   medicine and genetic analysis.
XX
SQ   Sequence 21 BP; 3 A; 6 C; 9 G; 3 T; 0 other;

  Query Match             82.0%;  Score 16.4;  DB 22;  Length 21;
  Best Local Similarity   94.4%;  Pred. No. 2.1e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 CCCCCAGGGCACTCATGG 20
           |||||||||| |||||||
Db      21 CCCCCAGGGCGCTCATGG 4


RESULT 3
AAZ31450
ID   AAZ31450 standard; DNA; 20 BP.
XX
AC   AAZ31450;
XX
DT   07-FEB-2000  (first entry)
XX
DE   Human neuropilin mRNA specific antisense oligo GTI3621.
XX
KW   Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW   cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9955855-A2.
XX
PD   04-NOV-1999.
XX
PF   23-APR-1999;   99WO-CA00324.
XX
PR   23-APR-1998;   98US-0082791.
XX
PA   (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI   Wright JA,  Young AH,  Lee YS;
XX
DR   WPI; 2000-023357/02.
XX
PT   Antisense oligonucleotides that inhibit neuropilin expression, useful
PT   for treating cancer -
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 4 A; 7 C; 5 G; 4 T; 0 other;

  Query Match             80.0%;  Score 16;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      5 CCCAGGGCACTCATGG 20
          ||||||||||||||||
Db      1 CCCAGGGCACTCATGG 16


RESULT 4
AAL27306
ID   AAL27306 standard; DNA; 51 BP.
XX
AC   AAL27306;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #514.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1536; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 10 A; 18 C; 11 G; 12 T; 0 other;

  Query Match             76.0%;  Score 15.2;  DB 22;  Length 51;
  Best Local Similarity   85.0%;  Pred. No. 7.9e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GACCCCCAGGGCACTCATGG 20
          || || ||||||| ||||||
Db     26 GAGCCTCAGGGCAGTCATGG 45


RESULT 5
ABN34554/c
ID   ABN34554 standard; DNA; 60 BP.
XX
AC   ABN34554;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:7302.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 7302; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 17 A; 14 C; 19 G; 10 T; 0 other;

  Query Match             76.0%;  Score 15.2;  DB 24;  Length 60;
  Best Local Similarity   85.0%;  Pred. No. 7.9e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GACCCCCAGGGCACTCATGG 20
          || || ||||||| ||||||
```

```
Db     46 GAGCCTCAGGGCAGTCATGG 27

RESULT 6
ABK88692/c
ID   ABK88692 standard; DNA; 39 BP.
XX
AC   ABK88692;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Human transmembrane activator and CAML-interactor (TACI), PCR primer #2.
XX
KW   Human; tumour necrosis factor; TNF delta; pulmonary system disorder;
KW   immunoglobulin production; B-cell proliferation; immune system disorder;
KW   autoimmune disease; cancer; lymphoproliferative disorder; pain;
KW   microbial infection; parasitic infection; bone disease; atherosclerosis;
KW   cardiovascular disorder; neurodegenerative disease; wound healing;
KW   graft versus host disease; haematopoietic cell disorder; nephritis;
KW   transmembrane activator and CAML-interactor; TACI; TNF epsilon; PCR;
KW   primer; ss.
XX
OS   Homo sapiens.
XX
PN   US2002064829-A1.
XX
PD   30-MAY-2002.
XX
PF   14-JUN-2001; 2001US-0879919.
XX
PR   14-MAR-1996;   96US-016812P.
PR   15-JUN-2000; 2000US-211537P.
PR   23-OCT-2000; 2000US-241952P.
PR   13-DEC-2000; 2000US-254875P.
PR   16-MAR-2001; 2001US-276248P.
PR   23-MAR-2001; 2001US-277978P.
PR   25-MAY-2001; 2001US-293499P.
PR   12-MAR-1997;   97US-0815783.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Yu G,  Ni J,  Gentz RL,  Dillon PJ;
XX
DR   WPI; 2002-556722/59.
XX
PT   Novel human multimeric tumour necrosis factor delta or epsilon protein
PT   useful for treating disease or disorder of immune system such as
PT   autoimmune disease, immunodeficiency, or cancer of immune system  -
XX
PS   Example 29; Page 115; 143pp; English.
XX
CC   The present invention relates to the isolation of human tumour necrosis
CC   factor (TNF) delta and TNF epsilon proteins, and the polynucleotide
CC   sequences encoding them. The proteins are useful for modulating
CC   immunoglobulin production or for modulating proliferation of B-cells.
CC   The sequences of the invention are useful for treating diseases or
CC   disorders of the immune system. Such disorders include autoimmune
CC   diseases (e.g. systemic lupus erythematosus (SLE), acquired
CC   immunodeficiency syndrome (AIDS)), cancers of the immune system
CC   (e.g. chronic lymphocytic leukaemia (CLL), multiple myeloma,
CC   non-Hodgkin's lymphoma or Hodgkin's disease), lymphoproliferative
CC   disorders, microbial infections (e.g. viral, bacterial), parasitic
CC   infections, nephritis, bone disease (e.g. osteoporosis), atherosclerosis,
CC   pain, cardiovascular disorders (e.g. myocardial infarction, stroke),
CC   neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
CC   disease), graft versus host disease, wound healing, haematopoietic cell
CC   disorders (e.g. anaemia), inflammatory disorders (e.g. asthma),
CC   diseases or disorders associated with various mucous membranes of the
CC   body (e.g. mucositis), and disorders of the pulmonary system. The
CC   proteins are also useful as a vaccine adjuvant that enhances immune
CC   responsiveness to specific antigens. The present sequence represents
CC   a PCR primer used to amplify human transmembrane activator and
CC   CAML-interactor (TACI) cDNA in the examples of the present invention.
XX
SQ   Sequence 39 BP; 8 A; 11 C; 13 G; 7 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 24;  Length 39;
  Best Local Similarity   93.8%;  Pred. No. 1.8e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 CCCAGGGCACTCATGG 20
          |||||| |||||||||
Db     32 CCCAGGCCACTCATGG 17


RESULT 7
ABK13410/c
ID   ABK13410 standard; DNA; 39 BP.
XX
AC   ABK13410;
XX
DT   09-APR-2002  (first entry)
XX
DE   Tumour necrosis factor (TNF) receptor TACI, 5' PCR primer #2.
XX
KW   Tumour necrosis factor; TNF; cytostatic; arteriosclerosis;
KW   analgesic; cerebroprotective; nootropic; neuroprotective; hepatotropic;
KW   immunoglobulin production; B cell proliferation; immunosuppressive;
KW   HIV; human immunodeficiency virus; autoimmune disease; immunodeficiency;
KW   Sjogren's syndrome; systemic lupus erythematosus; Hodgkin's disease;
KW   common variable immunodeficiency; CVID; non-Hodgkin's lymphoma; AIDS;
KW   acquired immunodeficiency virus; cancer; multiple myeloma; CLL;
KW   chronic lymphocytic leukaemia; lymphoproliferative disorder;
KW   bacterial infection; viral infection; osteoporosis; atherosclerosis;
KW   pain; cardiovascular disorder; stroke; allergy; Alzheimer's disease;
KW   neurodegenerative disease; inflammation; liver disease; cirrhosis;
KW   cardiomyopathy; diabetes; asthma; psoriasis; glomerulonephritis;
KW   ulcerative colitis; angiogenesis; septic shock; wound healing;
KW   PCR; primer; ss; tumour necrosis factor receptor; TACI.
XX
OS   Homo sapiens.
XX
PN   WO200196528-A2.
XX
PD   20-DEC-2001.
XX
PF   14-JUN-2001; 2001WO-US19026.
XX
PR   15-JUN-2000; 2000US-211537P.
PR   23-OCT-2000; 2000US-241952P.
PR   13-DEC-2000; 2000US-254875P.
PR   16-MAR-2001; 2001US-276248P.
PR   23-MAR-2001; 2001US-277978P.
PR   25-MAY-2001; 2001US-293499P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Yu G,  Ni J,  Gentz RL,  Dillon PJ,  Hilbert D;
XX
DR   WPI; 2002-130727/17.
XX
PT   Novel multimeric human tumour necrosis factor delta or epsilon protein
PT   useful for treating cancer, immune system disorders, infection,
PT   cardiovascular disorders, liver disease, cardiomyopathy, diabetes and
PT   psoriasis  -
XX
PS   Example 29; Page 295; 344pp; English.
XX
CC   The invention describes a multimeric human tumour necrosis factor (TNF)
CC   delta or epsilon protein (I). (I) or a composition containing them (II)
CC   are useful for modulating immunoglobulin production or proliferation of B
CC   cells. (I) or (II) is useful: for treating a disease or disorder of the
CC   immune system, preferably an autoimmune disease (e.g. Sjogren's syndrome,
CC   systemic lupus erythematosus or common variable immunodeficiency (CVID));
CC   an immunodeficiency e.g. acquired immunodeficiency syndrome (AIDS);
CC   cancer of the immune system (e.g. Hodgkin's disease, non-Hodgkin's
```

```
CC   lymphoma, multiple myeloma and chronic lymphocytic leukaemia (CLL)); in
CC   the diagnosis and treatment or prevention of cancer, lymphoproliferative
CC   disorder, bacterial and viral infections, osteoporosis, atherosclerosis,
CC   pain, cardiovascular disorders (e.g. stroke), allergy, inflammation,
CC   neurodegenerative disease (e.g. Alzheimer's disease), liver disease (e.g.
CC   cirrhosis), cardiomyopathy, diabetes, asthma, psoriasis, septic shock,
CC   glomerulonephritis, ulcerative colitis, arteriosclerosis; for promoting
CC   angiogenesis and wound healing; as a diagnostic research reagent; as an
CC   agent to target and kill cells expressing a TNFdelta and/or TNFepsilon
CC   receptor; in apoptosis of transformed cell lines; mediation of cell
CC   activation and proliferation; and as an immunogen to produce (II). (II)
CC   is useful to purify, detect and target (I), for measuring levels of (I)
CC   in biological samples, for immunophenotyping samples, and to treat,
CC   inhibit or prevent diseases and disorders associated with aberrant
CC   expression and/or activity of (I). This sequence represents a 5' PCR
CC   primer used to isolate the TNF receptor superfamily member TACI, required
CC   in an assay to establish if tumour necrosis factor epsilon (TNF-epsilon)
CC   binds the receptor described in the method of the invention.
XX
SQ   Sequence 39 BP; 8 A; 11 C; 13 G; 7 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 24;  Length 39;
  Best Local Similarity   93.8%;  Pred. No. 1.8e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CCCAGGGCACTCATGG 20
           |||||| |||||||||
Db      32 CCCAGGCCACTCATGG 17


RESULT 8
AAL27305
ID   AAL27305 standard; DNA; 51 BP.
XX
AC   AAL27305;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #513.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1536; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 12 A; 16 C; 11 G; 12 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ACCCCCAGGGCACTCATGG 20
             | || ||||||| ||||||
Db      30 AGCCTCAGGGCAGTCATGG 48


RESULT 9
AAL27307
ID   AAL27307 standard; DNA; 51 BP.
XX
AC   AAL27307;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #515.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1537; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
```

```
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 12 A; 16 C; 11 G; 12 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ACCCCCAGGGCACTCATGG 20
           | || ||||||| ||||||
Db       6 AGCCTCAGGGCAGTCATGG 24


RESULT 10
ABN59176/c
ID   ABN59176 standard; DNA; 60 BP.
XX
AC   ABN59176;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:31924.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 31924; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 15 A; 18 C; 13 G; 14 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 24;  Length 60;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           | |||| ||| ||||||||
Db      30 GTCCCCGAGGCCACTCATG 12


RESULT 11
AAL27308
ID   AAL27308 standard; DNA; 51 BP.
XX
AC   AAL27308;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #516.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1537; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
```

```
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 10 A; 16 C; 11 G; 14 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 3.5e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCCCAGGGCACTCATGG 20
           || ||||||| ||||||
Db       2 CCTCAGGGCAGTCATGG 18


RESULT 12
AAL28078/c
ID   AAL28078 standard; DNA; 51 BP.
XX
AC   AAL28078;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #1286.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1748; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 15 A; 13 C; 14 G; 9 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 22;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 4.3e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           | | ||||||||||| |||
Db      41 GTCTCCCAGGGCACTACTGG 22


RESULT 13
AAV18190
ID   AAV18190 standard; DNA; 59 BP.
XX
AC   AAV18190;
XX
DT   28-AUG-1998  (first entry)
XX
DE   Primer for Fanconi anaemia of complementation group A gene.
XX
KW   Fanconi anaemia of complementation group A; FA-A; genetic defect;
KW   prenatal FA-A; FA-A carrier detection; disease diagnosis; PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9814462-A1.
XX
PD   09-APR-1998.
XX
PF   03-OCT-1997;   97WO-US18010.
XX
PR   04-OCT-1996;   96US-0726012.
XX
PA   (FANC-) FANCONI ANEMIA RES FUND INC.
XX
PI   Joenje H,  Lo Ten Foe JR;
XX
DR   WPI; 1998-240012/21.
XX
PT   DNA for Fanconi Anaemia complementation group A - useful for, e.g.
PT   developing products for diagnosis and screening of disease and gene
PT   therapy
XX
PS   Disclosure; Page 10; 63pp; English.
XX
CC   This sequence represents a PCR primer for the DNA encoding the Fanconi
CC   anaemia of complementation group A (FA-A) protein of the invention. The
CC   amplified DNA's may be used to complement a genetic defect in a cell
CC   (especially the FA-A gene). The products can be used for screening
CC   (especially prenatal FA-A), detection of FA-A carriers and FA-A disease
CC   diagnosis.
XX
SQ   Sequence 59 BP; 18 A; 16 C; 12 G; 13 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 19;  Length 59;
  Best Local Similarity   80.0%;  Pred. No. 4.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           |||| ||| |||  ||||||
Db      27 GACCACCATGGCCATCATGG 46


RESULT 14
```

```
ABN39793/c
ID   ABN39793 standard; DNA; 60 BP.
XX
AC   ABN39793;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:12541.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 12541; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 14 A; 14 C; 18 G; 14 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 60;
  Best Local Similarity   80.0%;  Pred. No. 4.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           |||| ||||||||  || ||
Db      29 GACCTCCAGGGCATACAAGG 10


RESULT 15
```

```
AAF75272
ID   AAF75272 standard; DNA; 61 BP.
XX
AC   AAF75272;
XX
DT   09-MAY-2001  (first entry)
XX
DE   Human PPARgamma PCR primer #1.
XX
KW   Human; peroxisome proliferator-activator receptor delta; PPARdelta;
KW   RXR; cytostatic; nootropic; neuroprotective; anti-HIV; cardiant;
KW   cerebroprotective; vasotropic; antiulcer; immunosuppressive;
KW   nephrotropic; antibacterial; antiviral; antifungal; protozoacide;
KW   non-steroidal anti-inflammatory disease; NSAID; infection;
KW   Alzheimer's disease; AIDS; muscle wasting disease; autoimmune disease;
KW   PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200112858-A1.
XX
PD   22-FEB-2001.
XX
PF   16-AUG-2000; 2000WO-US22411.
XX
PR   16-AUG-1999;   99US-0148701.
PR   15-AUG-2000; 2000US-0148701.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   He T,  Kinzler KW,  Vogelstein B;
XX
DR   WPI; 2001-211236/21.
XX
PT   Novel subgenomic polynucleotide having peroxisome
PT   proliferator-activator receptor proliferator (PPAR-delta) and RXR
PT   binding elements used to identify downregulators of PPAR-delta
PT   transcriptional activity  -
XX
PS   Example 1; Page 21; 70pp; English.
XX
CC   The present sequence is given in a specification relating to an
CC   isolated subgenomic polynucleotide comprising a peroxisome
CC   proliferator-activator receptor (PPAR)delta binding element and an RXR
CC   binding element. The polynucleotide is useful for identifying potential
CC   therapeutic agents for cancer treatment and for ameliorating negative
CC   side effects of non-steroidal anti-inflammatory diseases (NSAIDs). Test
CC   compounds which increase transcription of PPARdelta protein, PPARdelta
CC   protein binding to a PPARdelta binding element, or expression of a
CC   reporter gene which is under the control of a PPARdelta binding element,
CC   are identified. These are candidates for use in encouraging cell
CC   proliferation or preventing cell apoptosis in a disease state such as
CC   Alzheimer's disease, AIDS, muscular dystrophy, amyotrophic lateral
CC   sclerosis, or other muscle wasting diseases, autoimmune diseases, heart
CC   attack, stroke, ischaemic heart disease, kidney failure, septic shock,
CC   or a disease in which the cell is infected with a pathogen, such as a
CC   virus, bacterium, fungus, mycoplasma, or protozoan, to promote healing
CC   of the stomach or intestines, or to ameliorate negative side effects of
CC   NSAIDs, such as gastric and intestinal ulceration.
XX
SQ   Sequence 61 BP; 20 A; 16 C; 14 G; 11 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 22;  Length 61;
  Best Local Similarity   80.0%;  Pred. No. 4.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           |||| ||| ||||  |||||
Db      28 GACCACCATGGCAACCATGG 47


Search completed: November 23, 2002, 06:29:03
Job time : 102.6 secs
```

THIS PAGE BLANK (USPTO)

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/067,615
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/596,387
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/382,454
;       FILING DATE:  01-FEB-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/283,302
;       FILING DATE:  29-JUL-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Pearson, Louise S.
;       REGISTRATION NUMBER:  32,369
;       REFERENCE/DOCKET NUMBER:  STR-4665-CIP2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (708) 267-5300
;       TELEFAX:  (708) 267-5376
;   INFORMATION FOR SEQ ID NO:  39:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  40 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
US-09-067-615-39

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 40;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           ||| ||||||||  | |||
Db      38 GACACCCAGGGCCTTTATG 20


RESULT 15
PCT-US95-09816A-39/c
; Sequence 39, Application PC/TUS9509816A
;  GENERAL INFORMATION:
;     APPLICANT:  Wong, Hing C.
;     APPLICANT:  Rhode, Peter R.
;     APPLICANT:  Widanz, Jon A.
;     APPLICANT:  Grammer, Susan
;     APPLICANT:  Edwards, Ana C.
;     APPLICANT:  Chavaillaz, Pierre-Andre
;     APPLICANT:  Jiao, Jin-An
;     TITLE OF INVENTION:  MHC COMPLEXES AND USES THEREOF
;     NUMBER OF SEQUENCES:  123
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Dade International, Inc.
;       STREET:  1717 Deerfield Road
;       CITY:  Deerfield
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60015
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/09816A
;       FILING DATE:  31-JUL-1995
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/382,454
;       FILING DATE:  01-FEB-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/283,302
;       FILING DATE:  29-JUL-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Pearson, Louise S.
;       REGISTRATION NUMBER:  32,369
;       REFERENCE/DOCKET NUMBER:  STR-4665-CIP2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (708) 267-5300
;       TELEFAX:  (708) 267-5376
;   INFORMATION FOR SEQ ID NO:  39:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  40 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
PCT-US95-09816A-39

  Query Match             63.0%;  Score 12.6;  DB 5;  Length 40;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           ||| ||||||||  | |||
Db      38 GACACCCAGGGCCTTTATG 20


Search completed: November 23, 2002, 06:36:13
Job time : 23.55 secs
```

```
  Query Match             63.0%;  Score 12.6;  DB 4;  Length 29;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           ||| |||||||| | |||
Db      29 GACACCCAGGGCCTTTATG 11


RESULT 12
US-08-670-175-7
; Sequence 7, Application US/08670175
; Patent No. 5854081
;  GENERAL INFORMATION:
;    APPLICANT:  LINDEN, JOEL
;    APPLICANT:  TAYLOR, HEIDI
;    APPLICANT:  ROBEVA, ANNA
;    APPLICANT:  WOODARD, ROBIN
;    APPLICANT:  JIN, XIAOWEI
;    TITLE OF INVENTION:  STABLE EXPRESSION OF HUMAN ADENOSINE
;    TITLE OF INVENTION:  RECEPTORS, AND ASSAYS EMPLOYING THE SAME
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/670,175
;      FILING DATE:  20-JUN-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  494-176-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "SYNTHETIC DNA PRIMER"
US-08-670-175-7

  Query Match             63.0%;  Score 12.6;  DB 2;  Length 30;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           || || |||| ||||| ||
Db       5 GATCCTCAGGACACTCCTG 23


RESULT 13
US-08-596-387B-39/c
; Sequence 39, Application US/08596387B
; Patent No. 5869270
;  GENERAL INFORMATION:
;    APPLICANT:  Rhode, Peter R.
;    APPLICANT:  Jiao, Jin-An
;    APPLICANT:  Burkhardt, Martin
;    APPLICANT:  Wong, Hing
;    TITLE OF INVENTION:  MHC COMPLEXES AND USES THEREOF
;    NUMBER OF SEQUENCES:  124
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dade International, Inc.
;      STREET:  1717 Deerfield Road
;      CITY:  Deerfield
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60015
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/596,387B
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09816
;      FILING DATE:  31-JUL-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/382,454
;      FILING DATE:  01-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/283,302
;      FILING DATE:  29-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pearson, Louise S.
;      REGISTRATION NUMBER:  32,369
;      REFERENCE/DOCKET NUMBER:  STR-4665-CIP2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (708) 267-5300
;      TELEFAX:  (708) 267-5376
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
US-08-596-387B-39

  Query Match             63.0%;  Score 12.6;  DB 2;  Length 40;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           ||| |||||||| | |||
Db      38 GACACCCAGGGCCTTTATG 20


RESULT 14
US-09-067-615-39/c
; Sequence 39, Application US/09067615
; Patent No. 6309645
;  GENERAL INFORMATION:
;    APPLICANT:  Rhode, Peter R.
;    APPLICANT:  Jiao, Jin-An
;    APPLICANT:  Burkhardt, Martin
;    APPLICANT:  Wong, Hing
;    TITLE OF INVENTION:  MHC COMPLEXES AND USES THEREOF
;    NUMBER OF SEQUENCES:  124
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dade International, Inc.
;      STREET:  1717 Deerfield Road
;      CITY:  Deerfield
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60015
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;      LOCATION:  1..69
US-08-352-179-10

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 69;
  Best Local Similarity   87.5%;  Pred. No. 1.5e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 CCCAGGGCACTCATGG 20
          |||||| | |||||||
Db     62 CCCAGGCCGCTCATGG 47


RESULT 9
US-08-617-874-4/c
; Sequence 4, Application US/08617874
; Patent No. 6338850
;  GENERAL INFORMATION:
;    APPLICANT:  Jevnikar, Anthony M.
;    APPLICANT:  Ma, Shengwu
;    APPLICANT:  Stiller, Calvin R.
;    TITLE OF INVENTION:  METHODS AND PRODUCTS FOR CONTROLLING
;    TITLE OF INVENTION:  IMMUNE RESPONSES IN MAMMALS
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis, L.L.P.
;      STREET:  George Mason Building, 699 Prince Street
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22314-3187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/617,874
;      FILING DATE:  21-MAY-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rea, Teresa Stanek
;      REGISTRATION NUMBER:  30,427
;      REFERENCE/DOCKET NUMBER:  024916-002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  28 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA - primer"
US-08-617-874-4

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 28;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GACCCCCAGGGCACTCATG 19
          ||| ||||||||  | |||
Db     27 GACACCCAGGGCCTTTATG 9


RESULT 10
US-08-617-874-6/c
; Sequence 6, Application US/08617874
; Patent No. 6338850
;  GENERAL INFORMATION:
;    APPLICANT:  Jevnikar, Anthony M.
;    APPLICANT:  Ma, Shengwu
;    APPLICANT:  Stiller, Calvin R.
;    TITLE OF INVENTION:  METHODS AND PRODUCTS FOR CONTROLLING
;    TITLE OF INVENTION:  IMMUNE RESPONSES IN MAMMALS
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis, L.L.P.
;      STREET:  George Mason Building, 699 Prince Street
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22314-3187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/617,874
;      FILING DATE:  21-MAY-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rea, Teresa Stanek
;      REGISTRATION NUMBER:  30,427
;      REFERENCE/DOCKET NUMBER:  024916-002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  28 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA - primer"
US-08-617-874-6

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 28;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GACCCCCAGGGCACTCATG 19
          ||| ||||||||  | |||
Db     27 GACACCCAGGGCCTTTATG 9


RESULT 11
US-09-194-285-4/c
; Sequence 4, Application US/09194285
; Patent No. 6355479
; GENERAL INFORMATION:
;  APPLICANT: Webb, Susan R.
;  APPLICANT:  Winqvist, Ola
;  APPLICANT:  Karlsson, Lars
;  APPLICANT:  Jackson, Michael R.
;  APPLICANT:  Peterson, Per A.
;  TITLE OF INVENTION: MHC Class II Antigen Presenting Systems
;  TITLE OF INVENTION:  and Methods for Activating CD4+ T Cells
;  FILE REFERENCE: TSRI 536.1
;  CURRENT APPLICATION NUMBER: US/09/194,285
;  CURRENT FILING DATE:  1999-04-12
;  PRIOR APPLICATION NUMBER: PCT/US97/08697
;  PRIOR FILING DATE: 1997-05-22
;  PRIOR APPLICATION NUMBER: US 60/018,175
;  PRIOR FILING DATE: 1996-05-23
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 29
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthesized
US-09-194-285-4
```

```
Qy       3 CCCCCAGGGCACTCATGG 20
           ||||| |||  ||||:||
Db      35 CCCCCGGGGAGCTCAUGG 52


RESULT 6
US-09-071-433-64
; Sequence 64, Application US/09071433A
; Patent No. 6197584
; GENERAL INFORMATION:
;  APPLICANT: Bennett, C. Frank
;  APPLICANT: Cowsert, Lex M
;  TITLE OF INVENTION: Antisense Modulation of CD40 Expression
;  FILE REFERENCE: RTS-0002
;  CURRENT APPLICATION NUMBER: US/09/071,433A
;  CURRENT FILING DATE: 1998-05-01
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
;   LENGTH: 18
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Synthetic
US-09-071-433-64

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 18;
  Best Local Similarity   87.5%;  Pred. No. 1.5e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 ACCCCCAGGGCACTCA 17
           ||| ||||||| ||||
Db       2 ACCACCAGGGCTCTCA 17


RESULT 7
US-07-728-220C-14/c
; Sequence 14, Application US/07728220C
; Patent No. 6020168
;  GENERAL INFORMATION:
;    APPLICANT:  MATSUO, HISAYUKI
;    APPLICANT:  KANGAWA, KENJI
;    APPLICANT:  TANAKA, SHOJI
;    APPLICANT:  FUCHIMURA, KAYOKO
;    APPLICANT:  TAWARAGI, YASUNORI
;    TITLE OF INVENTION:  PORCINE CNP GENE AND PRECURSOR PROTEIN
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN, DARBY & CUSHMAN
;      STREET:  ELEVENTH FLOOR, 1615 L STREET, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036-5601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/728,220C
;      FILING DATE:  19910712
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SCOTT, WATSON T.
;      REGISTRATION NUMBER:  26,581
;      REFERENCE/DOCKET NUMBER:  WTS/9437/91816
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)861-3000
;      TELEFAX:  (202)822-0944
;      TELEX:  671 4627 CUSH
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  66 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-07-728-220C-14

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 66;
  Best Local Similarity   87.5%;  Pred. No. 1.5e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 CCCAGGGCACTCATGG 20
           |||||| | |||||||
Db      62 CCCAGGCCGCTCATGG 47


RESULT 8
US-08-352-179-10/c
; Sequence 10, Application US/08352179
; Patent No. 5670340
;  GENERAL INFORMATION:
;    APPLICANT:  YABUTA, Masayuki
;    APPLICANT:  SUZUKI, Yuji
;    APPLICANT:  OHSUYE, Kazuhiro
;    APPLICANT:  OSHIMA, Takehiro
;    APPLICANT:  ONAI, Seiko
;    APPLICANT:  MAGOTA, Koji
;    APPLICANT:  TANAKA, Shoji
;    TITLE OF INVENTION:  PROCESS FOR PRODUCING PEPTIDE
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker and Mathis
;      STREET:  The George Mason Bldg., Washington & Prince
;      STREET:  Sts.
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/352,179
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/929,597
;      FILING DATE:  17-AUG-1992
;      APPLICATION NUMBER:  JP 3-320769
;      FILING DATE:  19-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 4-223520
;      FILING DATE:  31-JUL-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crane-Feury, Sharon E
;      REGISTRATION NUMBER:  36,113
;      REFERENCE/DOCKET NUMBER:  001560-175
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  plasmid pUCCNP1
;    FEATURE:
;      NAME/KEY:  CDS
```

```
RESULT 2
US-09-005-165-8/c
; Sequence 8, Application US/09005165
; Patent No. 6093542
;  GENERAL INFORMATION:
;    APPLICANT:  ROMANO, JOSEPH
;    APPLICANT:  SHURTLIFF, ROXANNE
;    APPLICANT:  PAL, RANAJIT
;    TITLE OF INVENTION:  ISOTHERMAL TRANSCRIPTION BASED
;    TITLE OF INVENTION:  AMPLIFICATION ASSAY FOR THE DETECTION AND QUANTITATION OF
;    TITLE OF INVENTION:  MACROPHAGE DERIVED CHEMOKINE RNA
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AKZO NOBEL PATNT DEPARTMENT
;      STREET:  1300 PICCARD DRIVE, SUITE 206
;      CITY:  ROCKVILLE
;      STATE:  MARYLAND
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/005,165
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KLESNER, SHARON N.
;      REGISTRATION NUMBER:  36,335
;      REFERENCE/DOCKET NUMBER:  MDC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-005-165-8

  Query Match             66.0%;  Score 13.2;  DB 3;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 9.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 CCCCCAGGGCACTCATGG 20
           | ||||||||||||  ||
Db      23 CACCCAGGGCACTCTGGG 6


RESULT 3
US-09-060-756-146
; Sequence 146, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
;  APPLICANT: Cole, Stewart
;  APPLICANT: Buchrieser-Brosch, Roland
;  APPLICANT: Gordon, Stephen
;  APPLICANT: Billault, Alain
;  TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
;  TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
;  TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
;  FILE REFERENCE: 3495-0169
;  CURRENT APPLICATION NUMBER: US/09/060,756
;  CURRENT FILING DATE: 1998-04-16
;  NUMBER OF SEQ ID NOS: 743
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 146
;   LENGTH: 75
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
US-09-060-756-146

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 75;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ACCCCCAGGGCACTCATG 19
           |||||   |||||||||
Db      30 ACCCCACAGGCACTCATG 47


RESULT 4
US-09-282-996-9/c
; Sequence 9, Application US/09282996
; Patent No. 6143502
; GENERAL INFORMATION:
;  APPLICANT: Grentzmann, Guido
;  APPLICANT: Gesteland, Raymond F.
;  APPLICANT: Atkins, John F.
;  TITLE OF INVENTION: Dual-Luciferase Reporter System
;  FILE REFERENCE: T5864.NP
;  CURRENT APPLICATION NUMBER: US/09/282,996
;  CURRENT FILING DATE: 1999-03-31
;  EARLIER APPLICATION NUMBER:
;  EARLIER FILING DATE:
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: WordPerfect 8.0
; SEQ ID NO 9
;   LENGTH: 80
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide luci2
US-09-282-996-9

  Query Match             66.0%;  Score 13.2;  DB 3;  Length 80;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 CCCCCAGGGCACTCATGG 20
           ||||| |||  |||||||
Db      50 CCCCCGGGGAGCTCATGG 33


RESULT 5
US-09-282-996-8
; Sequence 8, Application US/09282996
; Patent No. 6143502
; GENERAL INFORMATION:
;  APPLICANT: Grentzmann, Guido
;  APPLICANT: Gesteland, Raymond F.
;  APPLICANT: Atkins, John F.
;  TITLE OF INVENTION: Dual-Luciferase Reporter System
;  FILE REFERENCE: T5864.NP
;  CURRENT APPLICATION NUMBER: US/09/282,996
;  CURRENT FILING DATE: 1999-03-31
;  EARLIER APPLICATION NUMBER:
;  EARLIER FILING DATE:
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: WordPerfect 8.0
; SEQ ID NO 8
;   LENGTH: 82
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide luci1
US-09-282-996-8

  Query Match             66.0%;  Score 13.2;  DB 3;  Length 82;
  Best Local Similarity   77.8%;  Pred. No. 1e+03;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-11
Perfect score: 20
Sequence: 1 gacccccagggcactcatgg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13.6	68.0	59	2	US-08-726-012B-5	Sequence 5, Appli
c	2	13.2	66.0	24	3	US-09-005-165-8	Sequence 8, Appli
	3	13.2	66.0	75	4	US-09-060-756-146	Sequence 146, App
c	4	13.2	66.0	80	3	US-09-282-996-9	Sequence 9, Appli
	5	13.2	66.0	82	3	US-09-282-996-8	Sequence 8, Appli
	6	12.8	64.0	18	4	US-09-071-433-64	Sequence 64, Appl
c	7	12.8	64.0	66	3	US-07-728-220C-14	Sequence 14, Appl
c	8	12.8	64.0	69	1	US-08-352-179-10	Sequence 10, Appl
c	9	12.6	63.0	28	4	US-08-617-874-4	Sequence 4, Appli
c	10	12.6	63.0	28	4	US-08-617-874-6	Sequence 6, Appli
c	11	12.6	63.0	29	4	US-09-194-285-4	Sequence 4, Appli
	12	12.6	63.0	30	2	US-08-670-175-7	Sequence 7, Appli
c	13	12.6	63.0	40	2	US-08-596-387B-39	Sequence 39, Appl
c	14	12.6	63.0	40	4	US-09-067-615-39	Sequence 39, Appl
c	15	12.6	63.0	40	5	PCT-US95-09816A-39	Sequence 39, Appl
	16	12.6	63.0	45	2	US-08-596-387B-104	Sequence 104, App
	17	12.6	63.0	45	4	US-09-067-615-104	Sequence 104, App
	18	12.6	63.0	45	5	PCT-US95-09816A-104	Sequence 104, App
	19	12.4	62.0	18	1	US-08-164-200-14	Sequence 14, Appl
	20	12.4	62.0	24	2	US-08-288-508C-10	Sequence 10, Appl
	21	12.4	62.0	24	3	US-08-289-222E-20	Sequence 20, Appl
	22	12.4	62.0	24	4	US-09-054-526B-20	Sequence 20, Appl
	23	12.4	62.0	33	4	US-09-110-517-28	Sequence 28, Appl
c	24	12.4	62.0	37	4	US-09-263-904-7	Sequence 7, Appli
	25	12.4	62.0	38	2	US-08-292-620A-2308	Sequence 2308, Ap
	26	12.4	62.0	38	3	US-09-071-845-2308	Sequence 2308, Ap
c	27	12.2	61.0	22	5	PCT-US95-14418-38	Sequence 38, Appl
c	28	12.2	61.0	22	5	PCT-US95-15327-38	Sequence 38, Appl
	29	12.2	61.0	23	5	PCT-US95-14418-39	Sequence 39, Appl
	30	12.2	61.0	23	5	PCT-US95-15327-39	Sequence 39, Appl
c	31	12.2	61.0	24	1	US-07-961-268-1	Sequence 1, Appli
	32	12.2	61.0	30	2	US-08-816-605-5	Sequence 5, Appli
	33	12.2	61.0	31	3	US-08-946-914-55	Sequence 55, Appl
	34	12.2	61.0	31	4	US-09-656-450-55	Sequence 55, Appl
	35	12.2	61.0	32	3	US-08-946-914-57	Sequence 57, Appl
	36	12.2	61.0	32	4	US-09-656-450-57	Sequence 57, Appl
	37	12.2	61.0	33	4	US-09-527-236A-13	Sequence 13, Appl
	38	12.2	61.0	45	2	US-08-687-355A-14	Sequence 14, Appl
	39	12.2	61.0	45	4	US-09-407-367-14	Sequence 14, Appl
c	40	12.2	61.0	56	1	US-08-151-574-28	Sequence 28, Appl
c	41	12.2	61.0	56	2	US-08-419-448-28	Sequence 28, Appl
c	42	12.2	61.0	56	4	US-09-233-510-28	Sequence 28, Appl
	43	12.2	61.0	60	1	US-08-470-958-14	Sequence 14, Appl
c	44	12.2	61.0	80	1	US-08-471-985A-101	Sequence 101, App
c	45	12.2	61.0	80	5	PCT-US95-12401A-101	Sequence 101, App

ALIGNMENTS

```
RESULT 1
US-08-726-012B-5
; Sequence 5, Application US/08726012B
; Patent No. 5952190
;  GENERAL INFORMATION:
;    APPLICANT:  Hans Joenje, et al.
;    TITLE OF INVENTION:  cDNA FOR FANCONI ANEMIA COMPLEMENTATION GROUP A
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klarquist Sparkman Campbell Leigh & Whinston, LLP
;      STREET:  One World Trade Center, Suite 1600, 121 S.W. Salmon Street
;      CITY:  Portland
;      STATE:  OR
;      COUNTRY:  USA
;      ZIP:  97204-2988
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Disk, 3.5-inch
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  WordPerfect 5.1+, ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,012B
;      FILING DATE:  10/04/96
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Richard J. Polley
;      REGISTRATION NUMBER:  28,107
;      REFERENCE/DOCKET NUMBER:  3812-45520/RJP/DJE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (503) 226-7391
;      TELEFAX:  (503) 228-9446
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-726-012B-5

  Query Match             68.0%;  Score 13.6;  DB 2;  Length 59;
  Best Local Similarity   80.0%;  Pred. No. 6.6e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           |||| ||| |||  ||||||
Db      27 GACCACCATGGCCATCATGG 46
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-11
Perfect score: 20
Sequence: 1 gacccccagggcactcatgg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.4	72.0	39	10	US-09-879-919-16	Sequence 16, Appl
	2	13.4	67.0	78	10	US-09-938-700-25	Sequence 25, Appl
c	3	13.2	66.0	92	10	US-09-923-876-3869	Sequence 3869, Ap
	4	12.8	64.0	18	10	US-09-067-638B-64	Sequence 64, Appl
	5	12.8	64.0	25	10	US-09-305-856B-93	Sequence 93, Appl
c	6	12.8	64.0	88	10	US-09-864-761-21663	Sequence 21663, A
	7	12.8	64.0	94	10	US-09-864-761-21406	Sequence 21406, A
c	8	12.6	63.0	28	12	US-10-005-073-4	Sequence 4, Appli
c	9	12.6	63.0	28	12	US-10-005-073-6	Sequence 6, Appli
c	10	12.6	63.0	37	12	US-10-081-281-10	Sequence 10, Appl
c	11	12.6	63.0	40	10	US-09-848-164-39	Sequence 39, Appl
	12	12.6	63.0	45	10	US-09-848-164-104	Sequence 104, App
	13	12.6	63.0	79	10	US-09-864-761-24792	Sequence 24792, A
c	14	12.6	63.0	82	10	US-09-783-590-5203	Sequence 5203, Ap
c	15	12.6	63.0	92	10	US-09-864-761-23498	Sequence 23498, A
c	16	12.4	62.0	37	10	US-09-264-468B-16	Sequence 16, Appl
c	17	12.4	62.0	52	10	US-09-879-919-15	Sequence 15, Appl
	18	12.2	61.0	30	10	US-09-739-394-5	Sequence 5, Appli
	19	12.2	61.0	30	12	US-10-097-330-5	Sequence 5, Appli
	20	12.2	61.0	31	10	US-09-263-689-55	Sequence 55, Appl
	21	12.2	61.0	32	10	US-09-263-689-57	Sequence 57, Appl
	22	12.2	61.0	33	9	US-09-756-854-13	Sequence 13, Appl
	23	12.2	61.0	33	9	US-10-041-574-13	Sequence 13, Appl
c	24	12.2	61.0	67	10	US-09-050-516-1	Sequence 1, Appli
	25	12.2	61.0	72	10	US-09-918-063-62	Sequence 62, Appl
c	26	12.2	61.0	72	10	US-09-918-063-64	Sequence 64, Appl
c	27	12.2	61.0	80	9	US-09-736-457-660	Sequence 660, App
	28	12.2	61.0	91	10	US-09-983-965-5100	Sequence 5100, Ap
c	29	12.2	61.0	95	10	US-09-783-590-7186	Sequence 7186, Ap
	30	12.2	61.0	96	10	US-09-864-761-32598	Sequence 32598, A
	31	12.2	61.0	100	9	US-09-728-444-149	Sequence 149, App
c	32	12	60.0	27	10	US-09-799-160-5	Sequence 5, Appli
c	33	12	60.0	27	10	US-09-799-160-22	Sequence 22, Appl
c	34	12	60.0	30	10	US-09-799-160-4	Sequence 4, Appli
c	35	12	60.0	30	10	US-09-799-160-24	Sequence 24, Appl
	36	12	60.0	31	10	US-09-801-274-1599	Sequence 1599, Ap
c	37	12	60.0	31	10	US-09-799-160-18	Sequence 18, Appl
	38	12	60.0	39	10	US-09-473-872-32	Sequence 32, Appl
c	39	12	60.0	47	10	US-09-179-536B-255	Sequence 255, App
c	40	12	60.0	48	10	US-09-179-536B-256	Sequence 256, App
c	41	12	60.0	49	10	US-09-179-536B-257	Sequence 257, App
c	42	12	60.0	50	10	US-09-179-536B-258	Sequence 258, App
c	43	12	60.0	51	10	US-09-179-536B-259	Sequence 259, App
c	44	12	60.0	52	10	US-09-179-536B-260	Sequence 260, App
c	45	12	60.0	54	9	US-09-858-332-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-879-919-16/c
; Sequence 16, Application US/09879919
; Patent No. US20020064829A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Guo-Liang, et al.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Delta and Epsilon
; FILE REFERENCE: PF253P1
; CURRENT APPLICATION NUMBER: US/09/879,919
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,978
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/254,875
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/241,952
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/211,537
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 08/815,783
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: 60/016,812
; PRIOR FILING DATE: 1996-03-14
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-879-919-16

Query Match 72.0%; Score 14.4; DB 10; Length 39;
Best Local Similarity 93.8%; Pred. No. 2.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       5 CCCAGGGCACTCATGG 20
           |||||| |||||||||
Db      32 CCCAGGCCACTCATGG 17
```

```
RESULT 2
US-09-938-700-25
; Sequence 25, Application US/09938700
; Patent No. US20020064525A1
; GENERAL INFORMATION:
;  APPLICANT: Morsey, et al.
;  TITLE OF INVENTION: Anti-IgE Vaccines
;  FILE REFERENCE: PC10761A
;  CURRENT APPLICATION NUMBER: US/09/938,700
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
;   LENGTH: 78
;   TYPE: DNA
;   ORGANISM: DOG CH3/CH4 NUCLEOTIDE SEQUENCE
US-09-938-700-25

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 78;
  Best Local Similarity   93.3%;  Pred. No. 8.1e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CCCAGGGCACTCATG 19
           ||||||||| ||||||
Db      25 CCCAGGGCCCTCATG 39


RESULT 3
US-09-923-876-3869/c
; Sequence 3869, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
;  APPLICANT: Lalgudi, Raghunath V.
;  APPLICANT:  Kamigaki, Laura Y. (Ito)
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
;  FILE REFERENCE: PL-0012-1 CON
;  CURRENT APPLICATION NUMBER: US/09/923,876
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: 09/298,329
;  PRIOR FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: 60/085,331
;  PRIOR FILING DATE: 1998-05-05
;  NUMBER OF SEQ ID NOS: 6332
;  SOFTWARE: PERL Program
; SEQ ID NO 3869
;   LENGTH: 92
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020013958A1 700453962H1
;   NAME/KEY: unsure
;   LOCATION: 13, 20, 72
;   OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-3869

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 92;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 CCCCCAGGGCACTCATGG 20
           |||  | ||||||||||||
Db      71 CCCGAAAGGCACTCATGG 54


RESULT 4
US-09-067-638B-64
; Sequence 64, Application US/09067638B
; Patent No. US20020028923A1
;  GENERAL INFORMATION:
;    APPLICANT:  Lex M. Cowsert
;    APPLICANT:  Brenda F. Baker
;    APPLICANT:  John McNeil
;    APPLICANT:  Susan M. Freier
;    APPLICANT:  Henri M. Sasmor
;    APPLICANT:  Douglas G. Brooks
;    APPLICANT:  Cara Ohashi
;    APPLICANT:  Jacqueline R. Wyatt
;    APPLICANT:  Alexander Borchers
;    APPLICANT:  Timothy A. Vickers
;    TITLE OF INVENTION:  Identification of Genetic
;    TITLE OF INVENTION:  Targets for Modulation By Oligonucleotides and
;    TITLE OF INVENTION:  Generation of Oligonucleotides for Gene
;    TITLE OF INVENTION:  Modulation
;    NUMBER OF SEQUENCES:  112
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WOODCOCK WASHBURN KURTZ
;      ADDRESSEE:  MACKIEWICZ & NORRIS LLP
;      STREET:  1 LIBERTY PLACE 46TH FLOOR
;      CITY:  PHILADELPHIA
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 MB
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-Windows NT
;      SOFTWARE:  WORD PERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/067,638B
;      FILING DATE:  28-APR-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/081,483
;      FILING DATE:  13-APR-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  John W. Caldwell
;      REGISTRATION NUMBER:  28,937
;      REFERENCE/DOCKET NUMBER:  ISIS-2960
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  64:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-067-638B-64

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 18;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 ACCCCCAGGGCACTCA 17
           ||| ||||||| ||||
Db       2 ACCACCAGGGCTCTCA 17


RESULT 5
US-09-305-856B-93
; Sequence 93, Application US/09305856B
; Patent No. US20020061518A1
; GENERAL INFORMATION:
;  APPLICANT:  Penny, Laura
;  APPLICANT:  Galvin, Margaret
;  TITLE OF INVENTION:  Genotyping the Human
;  TITLE OF INVENTION:  UDP-Glucuronosyltransferase 1 (UGT1) Gene
;  FILE REFERENCE:  4389-7 (formerly SEQ-17CIP)
;  CURRENT APPLICATION NUMBER: US/09/305,856B
;  CURRENT FILING DATE:  1999-05-05
;  PRIOR APPLICATION NUMBER:  60/084,807
;  PRIOR FILING DATE:  1998-05-07
;  NUMBER OF SEQ ID NOS:  124
```

```
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 93
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-305-856B-93

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 1.5e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCCCAGGGCACTCATG 19
           ||||||| || |||||
Db       3 CCCCAGGCCAATCATG 18


RESULT 6
US-09-864-761-21663/c
; Sequence 21663, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21663
;   LENGTH: 88
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AF152364.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.99
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   4
;   OTHER INFORMATION: NT HIT: gi9506504, EVALUE 7.20e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BE970003.1, EVALUE 4.20e+00
US-09-864-761-21663

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 88;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 CCCAGGGCACTCATGG 20
           ||||||||||||  ||
Db      52 CCCAGGGCACTCCAGG 37


RESULT 7
US-09-864-761-21406
; Sequence 21406, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21406
;   LENGTH: 94
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   OTHER INFORMATION: MAP TO D87009.1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
;   OTHER INFORMATION: SWISSPROT HIT: P04211, EVALUE 1.00e-07
;   OTHER INFORMATION: EST_HUMAN HIT: AW407157.1, EVALUE 2.00e-24
;   OTHER INFORMATION: NT HIT: Z73660.1, EVALUE 3.00e-46
US-09-864-761-21406

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 94;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CCCCCAGGGCACTCAT 18
           |||||||| |||| ||
Db      62 CCCCCAGGACACTGAT 77


RESULT 8
US-10-005-073-4/c
; Sequence 4, Application US/10005073
; Patent No. US20020090371A1
;   GENERAL INFORMATION:
;        APPLICANT: Jevnikar, Anthony M.
;                   Ma, Shengwu
;                   Stiller, Calvin R.
;        TITLE OF INVENTION: METHODS AND PRODUCTS FOR CONTROLLING
;                            IMMUNE RESPONSES IN MAMMALS
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Burns, Doane, Swecker & Mathis, L.L.P.
;             STREET: George Mason Building, 699 Prince Street
;             CITY: Alexandria
;             STATE: Virginia
;             COUNTRY: USA
;             ZIP: 22314-3187
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/005,073
;             FILING DATE: 07-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/617,874
;             FILING DATE: 21-MAY-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Rea, Teresa Stanek
;             REGISTRATION NUMBER: 30,427
;             REFERENCE/DOCKET NUMBER: 024916-002
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (703) 836-6620
;             TELEFAX: (703) 836-2021
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 28 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: other nucleic acid
;             DESCRIPTION: /desc = "DNA - primer"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-005-073-4

  Query Match             63.0%;  Score 12.6;  DB 12;  Length 28;
  Best Local Similarity   78.9%;  Pred. No. 1.8e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           ||| |||||||| | |||
Db      27 GACACCCAGGGCCTTTATG 9


RESULT 9
US-10-005-073-6/c
; Sequence 6, Application US/10005073
; Patent No. US20020090371A1
;   GENERAL INFORMATION:
;        APPLICANT: Jevnikar, Anthony M.
;                   Ma, Shengwu
;                   Stiller, Calvin R.
;        TITLE OF INVENTION: METHODS AND PRODUCTS FOR CONTROLLING
;                            IMMUNE RESPONSES IN MAMMALS
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Burns, Doane, Swecker & Mathis, L.L.P.
;             STREET: George Mason Building, 699 Prince Street
;             CITY: Alexandria
;             STATE: Virginia
;             COUNTRY: USA
;             ZIP: 22314-3187
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/005,073
;             FILING DATE: 07-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/617,874
;             FILING DATE: 21-MAY-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Rea, Teresa Stanek
;             REGISTRATION NUMBER: 30,427
;             REFERENCE/DOCKET NUMBER: 024916-002
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (703) 836-6620
;             TELEFAX: (703) 836-2021
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 28 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: other nucleic acid
;             DESCRIPTION: /desc = "DNA - primer"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-005-073-6

  Query Match             63.0%;  Score 12.6;  DB 12;  Length 28;
  Best Local Similarity   78.9%;  Pred. No. 1.8e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           ||| |||||||| | |||
Db      27 GACACCCAGGGCCTTTATG 9


RESULT 10
US-10-081-281-10/c
; Sequence 10, Application US/10081281
; Patent No. US20020151707A1
;   GENERAL INFORMATION:
;        APPLICANT: Kindsvogel, Wayne
;                   Gross, Jane A.
;                   Sheppard, Paul
```

```
;       TITLE OF INVENTION: Immune Mediators and Related Methods
;       NUMBER OF SEQUENCES: 121
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Townsend and Townsend and Crew LLP
;            STREET: Two Embarcadero Center, Eighth Floor
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94111-3834
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/081,281
;            FILING DATE: 20-Feb-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/261,811A
;            FILING DATE: 03-Mar-1999
;            APPLICATION NUMBER: US 08/480,002
;            FILING DATE: 07-JUN-1995
;            APPLICATION NUMBER: US 08/482,133
;            FILING DATE: 07-JUN-1995
;            APPLICATION NUMBER: US 08/483,241
;            FILING DATE: 07-JUN-1995
;            APPLICATION NUMBER: US 60/005,964
;            FILING DATE: 27-OCT-1995
;            APPLICATION NUMBER: US 08/657,581
;            FILING DATE: 07-JUN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Parent, Annette S.
;            REGISTRATION NUMBER: 42,058
;            REFERENCE/DOCKET NUMBER: 014058-005630US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 576-0200
;            TELEFAX: (415) 576-0300
;   INFORMATION FOR SEQ ID NO: 10:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 37 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-081-281-10

  Query Match             63.0%;  Score 12.6;  DB 12;  Length 37;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           ||| ||||||||  | |||
Db      32 GACACCCAGGGCCTTTATG 14


RESULT 11
US-09-848-164-39/c
; Sequence 39, Application US/09848164
; Patent No. US20020034513A1
;   GENERAL INFORMATION:
;       APPLICANT: Rhode, Peter R.
;                  Jiao, Jin-An
;                  Burkhardt, Martin
;                  Wong, Hing
;       TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
;       NUMBER OF SEQUENCES: 124
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Dade International, Inc.
;            STREET: 1717 Deerfield Road
;            CITY: Deerfield
;            STATE: Illinois
;            COUNTRY: USA
;            ZIP: 60015
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/848,164
;            FILING DATE: 03-May-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/067,615
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: US 08/382,454
;            FILING DATE: 01-FEB-1995
;            APPLICATION NUMBER: US 08/283,302
;            FILING DATE: 29-JUL-1994
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Pearson, Louise S.
;            REGISTRATION NUMBER: 32,369
;            REFERENCE/DOCKET NUMBER: STR-4665-CIP2
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (708) 267-5300
;            TELEFAX: (708) 267-5376
;   INFORMATION FOR SEQ ID NO: 39:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 40 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: unknown
;            TOPOLOGY: unknown
;       SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-848-164-39

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 40;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           ||| ||||||||  | |||
Db      38 GACACCCAGGGCCTTTATG 20


RESULT 12
US-09-848-164-104
; Sequence 104, Application US/09848164
; Patent No. US20020034513A1
;   GENERAL INFORMATION:
;       APPLICANT: Rhode, Peter R.
;                  Jiao, Jin-An
;                  Burkhardt, Martin
;                  Wong, Hing
;       TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
;       NUMBER OF SEQUENCES: 124
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Dade International, Inc.
;            STREET: 1717 Deerfield Road
;            CITY: Deerfield
;            STATE: Illinois
;            COUNTRY: USA
;            ZIP: 60015
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/848,164
;            FILING DATE: 03-May-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/067,615
;            FILING DATE: <Unknown>
```

```
;             APPLICATION NUMBER: US 08/382,454
;             FILING DATE: 01-FEB-1995
;             APPLICATION NUMBER: US 08/283,302
;             FILING DATE: 29-JUL-1994
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Pearson, Louise S.
;             REGISTRATION NUMBER: 32,369
;             REFERENCE/DOCKET NUMBER: STR-4665-CIP2
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (708) 267-5300
;             TELEFAX: (708) 267-5376
;   INFORMATION FOR SEQ ID NO: 104:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 45 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: unknown
;             TOPOLOGY: unknown
;        SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-09-848-164-104

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 45;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           ||| ||||||||  | |||
Db      20 GACACCCAGGGCCTTTATG 38


RESULT 13
US-09-864-761-24792
; Sequence 24792, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24792
;   LENGTH: 79
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC011483.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =   3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.1
;   OTHER INFORMATION: SWISSPROT HIT: P48046, EVALUE 9.60e+00
;   OTHER INFORMATION: NT HIT: U37672.1, EVALUE 3.00e-12
US-09-864-761-24792

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 79;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           |||  ||||||| || |||
Db      53 GACTACCAGGGCCCTGATG 71


RESULT 14
US-09-783-590-5203/c
; Sequence 5203, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
;  APPLICANT: Dillon, Patrick J.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Li, Haodong
;  APPLICANT:  Rosen, Craig A.
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
;  FILE REFERENCE: PO-16.2C1
;  CURRENT APPLICATION NUMBER: US/09/783,590
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 08/420,856
;  PRIOR FILING DATE: 1995-04-12
;  PRIOR APPLICATION NUMBER: 08/346,731
;  PRIOR FILING DATE: 1994-11-21
;  NUMBER OF SEQ ID NOS: 12485
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5203
;   LENGTH: 82
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (45)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-5203

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 82;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATG 19
           | || ||||||  ||||||
Db      21 GGCCTCCAGGGGGCTCATG 3
```

```
RESULT 15
US-09-864-761-23498/c
; Sequence 23498, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23498
;   LENGTH: 92
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL139352.8
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
;   OTHER INFORMATION: EST_HUMAN HIT: BE697302.1, EVALUE 1.80e-02
;   OTHER INFORMATION: SWISSPROT HIT: P36130, EVALUE 9.50e+00
;   OTHER INFORMATION: NT HIT: AF193612.1, EVALUE 1.20e-02
US-09-864-761-23498

  Query Match           63.0%;  Score 12.6;  DB 10;  Length 92;
  Best Local Similarity  78.9%;  Pred. No. 2e+03;
    Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ACCCCCAGGGCACTCATGG 20
           |||||| ||| || | |||
Db      62 ACCCCCTGGGGACCCCTGG 44


Search completed: November 23, 2002, 06:42:09
Job time : 18.25 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-11
Perfect score: 20
Sequence: 1 gacccccagggcactcatgg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	14.8	74.0	24	17	AZ428700	AZ428700 1M0212L16
c	2	14.2	71.0	76	9	AI304064	AI304064 ui63g07.y
	3	13.8	69.0	100	9	AL595750	AL595750 AL595750
c	4	13.6	68.0	50	9	AU105199	AU105199 AU105199
c	5	13.6	68.0	50	9	AU105201	AU105201 AU105201
c	6	13.6	68.0	50	9	AU105204	AU105204 AU105204
c	7	13.6	68.0	50	9	AU105205	AU105205 AU105205
c	8	13.6	68.0	50	9	AU105206	AU105206 AU105206
c	9	13.6	68.0	50	9	AU105207	AU105207 AU105207
	10	13.6	68.0	57	9	AU259073	AU259073 AU259073
	11	13.6	68.0	70	9	AI829164	AI829164 wk76a07.x
	12	13.4	67.0	70	9	AL800251	AL800251 AL800251
c	13	13.2	66.0	66	14	BQ248668	BQ248668 TaE25005H
	14	13.2	66.0	92	9	AA525968	AA525968 ni93g06.s
c	15	13.2	66.0	97	9	AI756476	AI756476 EtESTea15
	16	12.8	64.0	52	9	AI310821	AI310821 qo93g12.x
c	17	12.8	64.0	52	14	W90057	W90057 zh69h10.r1
	18	12.8	64.0	57	17	B01045	B01045 cSRL-123g8-
	19	12.8	64.0	57	17	B06891	B06891 cSRL-88A11-
c	20	12.8	64.0	58	17	AZ767622	AZ767622 1M0566J21
c	21	12.8	64.0	78	9	AA958370	AA958370 ua10f08.r
	22	12.6	63.0	35	14	N94609	N94609 zb79d03.s1
c	23	12.6	63.0	38	17	BH751579	BH751579 SALK_0503
c	24	12.6	63.0	45	17	BH619149	BH619149 SALK_0402
	25	12.6	63.0	46	9	AI744340	AI744340 tr09g03.x
c	26	12.6	63.0	46	17	BH753339	BH753339 SALK_0287
c	27	12.6	63.0	49	17	BH171096	BH171096 SALK_0037
c	28	12.6	63.0	49	17	BH213409	BH213409 SALK_0091
c	29	12.6	63.0	62	9	AA489456	AA489456 ae31a12.s
c	30	12.6	63.0	62	17	AZ819711	AZ819711 2M0091B05
c	31	12.6	63.0	63	17	BH617146	BH617146 SALK_0360
c	32	12.6	63.0	71	14	F37653	F37653 HSPD05621 H
	33	12.6	63.0	71	17	AZ308894	AZ308894 1M0012D01
c	34	12.6	63.0	72	17	BH750287	BH750287 SALK_0376
c	35	12.6	63.0	85	9	AA245159	AA245159 mx88h03.r
	36	12.6	63.0	86	9	AI783748	AI783748 tu45c07.x
c	37	12.6	63.0	87	17	BH633158	BH633158 SALK_0403
c	38	12.6	63.0	88	17	BH804906	BH804906 1008106A0
c	39	12.6	63.0	91	9	AA559317	AA559317 nl70a12.s
c	40	12.6	63.0	93	9	AA623328	AA623328 vi19g10.r
c	41	12.6	63.0	94	10	BE059176	BE059176 sn27c04.y
c	42	12.6	63.0	95	17	CNS03IGR	AL245556 Tetraodon
c	43	12.6	63.0	98	12	BG793956	BG793956 UTSW_SM18
c	44	12.6	63.0	98	17	BH750078	BH750078 SALK_0346
c	45	12.6	63.0	100	9	AA779639	AA779639 af31e04.s

ALIGNMENTS

RESULT 1
AZ428700/c
LOCUS AZ428700 24 bp DNA linear GSS 03-OCT-2000
DEFINITION 1M0212L16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0212L16 F, DNA sequence.
ACCESSION AZ428700
VERSION AZ428700.1 GI:10552713
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 24)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

```
            Plate: 0212  row: L  column: 16
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 24.
FEATURES              Location/Qualifiers
     source           1. .24
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0212L16"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT        4 a      6 c     10 g      4 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 17;  Length 24;
  Best Local Similarity   88.9%;  Pred. No. 1e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CCCCCAGGGCACTCATGG 20
           ||||| | ||||||||||
Db      18 CCCCCTGAGCACTCATGG 1


RESULT 2
AI304064/c
LOCUS       AI304064                 76 bp    mRNA    linear   EST 08-DEC-1998
DEFINITION  ui63g07.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1887132 5' similar to gb:M16961 ALPHA-2-HS-GLYCOPROTEIN
            PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AI304064
VERSION     AI304064.1  GI:3987814
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:971456
            Trace considered overall poor quality
            Seq primer: custom primer used
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .76
                      /organism="Mus musculus"
                      /strain="C57BL"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:1887132"
                      /clone_lib="Sugano mouse liver mlia"
                      /sex="female"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                      (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                      was primed with an oligo(dT) primer
                      [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                      ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                      and cloned into distinct DraIII sites of the pME18S-FL3
                      vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                      be used to isolate the cDNA insert.  Size selection was
                      performed to exclude fragments <1.5kb.  Library
                      constructed by Dr. Sumio Sugano (University of Tokyo
                      Institute of Medical Science).  Custom primers  for
                      sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                      primer CGACCTGCAGCTCGAGCACA."
BASE COUNT       12 a     25 c     20 g     19 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 76;
  Best Local Similarity   84.2%;  Pred. No. 2.6e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ACCCCCAGGGCACTCATGG 20
           |||||||||| ||||||
Db      31 ACCCCCAGGGATGTCATGG 13


RESULT 3
AL595750
LOCUS       AL595750                100 bp    mRNA    linear   EST 30-JUL-2001
DEFINITION  AL595750 XGC-gastrula Silurana tropicalis cDNA clone TGas004g01 5',
            mRNA sequence.
ACCESSION   AL595750
VERSION     AL595750.1  GI:15007825
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Huckle,E., Taylor,R., McMurray,A., Ashurst,J.L., Zorn,A.M. and
            Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Huckle E
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TGas004g01.sp6
            Sequencing primer: SP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
FEATURES              Location/Qualifiers
     source           1. .100
                      /organism="Silurana tropicalis"
                      /db_xref="taxon:8364"
                      /clone="TGas004g01"
                      /clone_lib="XGC-gastrula"
```

```
                     /dev_stage="gastrula (stages 10.5-13 mixed)"
                     /lab_host="Escherichia coli XL1-blue"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from stages
                     10-13 gastrulae.  EcoRI-NotI cut cDNA was then ligated
                     into pCS107 with EcoRI at the 5' end and NotI at the 3'
                     end."
BASE COUNT       19 a     24 c     34 g     23 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 100;
  Best Local Similarity   88.2%;  Pred. No. 4.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GACCCCCAGGGCACTCA 17
          | ||||| |||||||||
Db     41 GGCCCCCGGGGCACTCA 57


RESULT 4
AU105199/c
LOCUS       AU105199                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105199 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC00642, mRNA sequence.
ACCESSION   AU105199
VERSION     AU105199.1  GI:13554720
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HRC00642"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        5 a     21 c     16 g      8 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 4.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GACCCCCAGGGCACTCATGG 20
          | || ||||||| | |||||
Db     35 GGCCTCCAGGGCTCCCATGG 16


RESULT 5
AU105201/c
LOCUS       AU105201                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105201 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC04501, mRNA sequence.
ACCESSION   AU105201
VERSION     AU105201.1  GI:13554722
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HRC04501"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        6 a     20 c     16 g      8 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 4.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GACCCCCAGGGCACTCATGG 20
          | || ||||||| | |||||
Db     41 GGCCTCCAGGGCTCCCATGG 22


RESULT 6
AU105204/c
LOCUS       AU105204                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105204 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC06024, mRNA sequence.
ACCESSION   AU105204
VERSION     AU105204.1  GI:13554725
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
```

```
                     /db_xref="taxon:9606"
                     /clone="HRC06024"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT         4 a     22 c     16 g      8 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 4.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           | || ||||||| | |||||
Db      34 GGCCTCCAGGGCTCCCATGG 15


RESULT 7
AU105205/c
LOCUS       AU105205                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105205 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC07843, mRNA sequence.
ACCESSION   AU105205
VERSION     AU105205.1  GI:13554726
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HRC07843"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT         5 a     21 c     16 g      8 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 4.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           | || ||||||| | |||||
Db      35 GGCCTCCAGGGCTCCCATGG 16


RESULT 8
AU105206/c
LOCUS       AU105206                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105206 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC08436, mRNA sequence.
ACCESSION   AU105206
VERSION     AU105206.1  GI:13554727
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HRC08436"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT         6 a     18 c     17 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 4.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           | || ||||||| | |||||
Db      41 GGCCTCCAGGGCTCCCATGG 22


RESULT 9
AU105207/c
LOCUS       AU105207                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105207 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            KAT01893, mRNA sequence.
ACCESSION   AU105207
VERSION     AU105207.1  GI:13554728
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="KAT01893"
```

```
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        5 a      21 c      16 g       8 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 4.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           | || ||||||| | |||||
Db      36 GGCCTCCAGGGCTCCCATGG 17


RESULT 10
AU259073
LOCUS       AU259073                 57 bp    mRNA    linear   EST 25-APR-2002
DEFINITION  AU259073 3'-directed mouse cDNA library Mus musculus cDNA clone
            BED0014473 3', mRNA sequence.
ACCESSION   AU259073
VERSION     AU259073.1  GI:20325251
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Kato,K. and Matoba,R.
  TITLE     Generation of expressed sequence tags from mouse brain
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kikuya Kato
            Graduate School of Biological Sciences
            Nara Institute of Science and Technology
            8916-5 Takayama, Ikoma, Nara 630-0101, Japan
            Tel: 81-743-72-5581
            Fax: 81-743-72-5589
            Email: kkato@bs.aist-nara.ac.jp,
            URL:http://love2.aist-nara.ac.jp/BED/index.html.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="BED0014473"
                     /clone_lib="3'-directed mouse cDNA library"
                     /tissue_type="brain"
                     /note="Vector: pGEM-T-easy"
BASE COUNT       23 a      16 c      11 g       7 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 57;
  Best Local Similarity   80.0%;  Pred. No. 4.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           || ||||||| | | |||||
Db       1 GATCCCCAGGACCCACATGG 20


RESULT 11
AI829164
LOCUS       AI829164                 70 bp    mRNA    linear   EST 07-MAR-2000
DEFINITION  wk76a07.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2421300 3'
            similar to SW:CA12_MOUSE P28481 PROCOLLAGEN ALPHA 1(II) CHAIN
            PRECURSOR [CONTAINS: CHONDROCALCIN]. [3] TR:Q62031 TR:Q62032
            ;contains element MSR1 repetitive element ;, mRNA sequence.
ACCESSION   AI829164
VERSION     AI829164.1  GI:5449835
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11548-013
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1747   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2421300"
                     /clone_lib="NCI_CGAP_Pan1"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.72 kb. Life Technologies catalog #:
                     11548-013"
BASE COUNT       11 a      32 c      23 g       4 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 70;
  Best Local Similarity   80.0%;  Pred. No. 4.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCATGG 20
           ||||||| |||| | || ||
Db      27 GACCCCCGGGGCCCCCAGGG 46


RESULT 12
AL800251
LOCUS       AL800251                 70 bp    mRNA    linear   EST 27-JUN-2002
DEFINITION  AL800251 XGC-neurula Silurana tropicalis cDNA clone TNeu067i19 5',
            mRNA sequence.
ACCESSION   AL800251
VERSION     AL800251.1  GI:21586312
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2002
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Taylor R
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TNeu067i19.p1caSP6
            Sequencing primer: P1CSP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="TNeu067i19"
                     /clone_lib="XGC-neurula"
```

```
                     /dev_stage="neurula"
                     /lab_host="Escherichia coli DH10B"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from neurula.
                     EcoRI-NotI cut cDNA was then ligated into pCS107 with
                     EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT       12 a     29 c     18 g     11 t
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 9;  Length 70;
  Best Local Similarity  93.3%;  Pred. No. 5.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GACCCCCAGGGCACT 15
          | ||||||||||||||
Db     56 GCCCCCCAGGGCACT 70


RESULT 13
BQ248668/c
LOCUS       BQ248668                 66 bp    mRNA    linear   EST 03-MAY-2002
DEFINITION  TaE25005H12F TaE25 Triticum aestivum cDNA clone TaE25005H12F, mRNA
            sequence.
ACCESSION   BQ248668
VERSION     BQ248668.1  GI:20444544
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Cloutier,S.
  TITLE     Wheat functional genomics - Glenlea developing seeds cDNA libraries
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Dr. Sylvie Cloutier
            Cereal Research Centre, Agriculture and Agri-food Canada
            195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
            Tel: (204) 983-2340
            Fax: (204) 983-4604
            Email: scloutier@em.agr.ca
            was cloned directionally, not all sequences generated with reverse
            primer were from the 5' end (same with forward primer and 3' end).
            Average insert size is >870 bp
            Plate: 005  row: H  column: 12
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Triticum aestivum"
                     /cultivar="Glenlea"
                     /db_xref="taxon:4565"
                     /clone="TaE25005H12F"
                     /clone_lib="TaE25"
                     /tissue_type="developing seeds"
                     /dev_stage="25 days after anthesis"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pCMV-SPORT6.0 (invitrogen Technologies);
                     Site_1: NotI; Site_2: MluI; mRNA obtained from wheat seeds
                     of cultivar Glenlea 25 days post-anthesis"
BASE COUNT       13 a     17 c     19 g     17 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 14;  Length 66;
  Best Local Similarity  83.3%;  Pred. No. 6.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ACCCCCAGGGCACTCATG 19
          |||||||||| |  |||
Db     26 ACCCCCAGGGCCCCAATG 9


RESULT 14
AA525968
LOCUS       AA525968                 92 bp    mRNA    linear   EST 05-AUG-1997
DEFINITION  ni93g06.s1 NCI_CGAP_Pr21 Homo sapiens cDNA clone IMAGE:984442 3'
            similar to gb:M33882 INTERFERON-REGULATED RESISTANCE GTP-BINDING
            PROTEIN (HUMAN);, mRNA sequence.
ACCESSION   AA525968
VERSION     AA525968.1  GI:2268037
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 240   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:984442"
                     /clone_lib="NCI_CGAP_Pr21"
                     /sex="male"
                     /tissue_type="normal prostate"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; 1st strand cDNA was prepared
                     from normal prostate bulk tissue, and was then primed with
                     a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is not normalized. Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       30 a     19 c     26 g     17 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 9;  Length 92;
  Best Local Similarity  83.3%;  Pred. No. 7.4e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ACCCCCAGGGCACTCATG 19
          || | || ||||||||||
Db     50 ACACACACGGCACTCATG 67

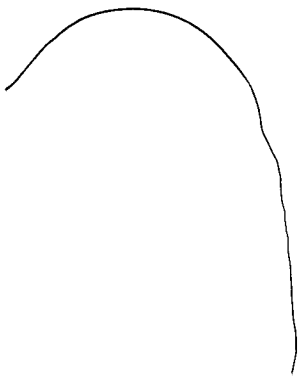

RESULT 15
AI756476/c
LOCUS       AI756476                 97 bp    mRNA    linear   EST 18-JAN-2000
DEFINITION  EtESTeal5d11.y1 Eimeria M5-6 Merozoite stage Eimeria tenella cDNA
            5' similar to SW:RLA2_BABBO P27055 60S ACIDIC RIBOSOMAL PROTEIN P2
            ;, mRNA sequence.
ACCESSION   AI756476
VERSION     AI756476.1  GI:5150199
KEYWORDS    EST.
SOURCE      Eimeria tenella.
  ORGANISM  Eimeria tenella
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
            Eimeria.
```

```
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Liberator,P., Diaz,C., Tang,K., Marra,M., Hillier,L., Kucaba,T.,
            Martin,J., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen
            ,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey
            ,N., Schurk,R., Ritter,E., Kohn,S., Florence,N., Shin,T., Jackson
            ,Y., Cardenas,M., McCann,R., Waterston,R., Wilson,R. and Sibley,D.
  TITLE     WashU-Merck Eimeria tenella project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Sibley, Ph.D.
            WashU-Merck Eimeria tenella project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Contact David Sibley (toxoest@borcim.wustl.edu) for further
            information relating to organism, libraries, or clone availability.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40RP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="Eimeria tenella"
                     /strain="LS18"
                     /db_xref="taxon:5802"
                     /clone_lib="Eimeria M5-6 Merozoite stage"
                     /dev_stage="Merozoite"
                     /lab_host="SOLR E. coli"
                     /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI
                     ; Merozoites were obtained from ceacal scrapings of
                     chickens infected with E. tenella.  The library may
                     contain a small percentage of host or bacterial
                     contaminents.  cDNA was synthesized from poly mRNA using
                     an oligo-dT primer containing a XhoI site.  Following
                     second strand synthesis, EcoRI adapters were ligated to
                     the cDNA and products were size-selected on Sephacryl
                     S500.  cDNAs were digested with EcoRI/XhoI and cloned into
                     lambda Zap II (Stratagene).  Clones were converted to
                     phagemids by mass excision using ExAssist helper phage and
                     SOLR cells (Stratagene ).  Insert sizes range from 0.7-1.5
                     kb."
BASE COUNT       17 a     19 c     31 g     30 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 97;
  Best Local Similarity   83.3%;  Pred. No. 7.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GACCCCCAGGGCACTCAT 18
           | |||||||| || ||||
Db      46 GCCCCCCAGGACAGTCAT 29


Search completed: November 26, 2002, 04:08:02
Job time : 765.8 secs
```



THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-12
Perfect score: 20
Sequence: 1 cgaccccacagacagccccc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		14.2	71.0	60	6	AR035484	AR035484 Sequence
2	c	13.8	69.0	25	6	AR003361	AR003361 Sequence
3	c	13.6	68.0	26	6	AR168794	AR168794 Sequence
4	c	13.6	68.0	26	6	AR168799	AR168799 Sequence
5	c	13.6	68.0	26	6	AR168801	AR168801 Sequence
6	c	13.6	68.0	26	6	AR200263	AR200263 Sequence
7	c	13.6	68.0	26	6	AR200268	AR200268 Sequence
8	c	13.6	68.0	26	6	AR200270	AR200270 Sequence
9	c	13.6	68.0	26	6	I27788	I27788 Sequence 20
10	c	13.6	68.0	26	6	I27793	I27793 Sequence 25
11	c	13.6	68.0	26	6	I27795	I27795 Sequence 27
12	c	13.6	68.0	45	6	AR168795	AR168795 Sequence
13	c	13.6	68.0	45	6	AR168796	AR168796 Sequence
14	c	13.6	68.0	45	6	AR200264	AR200264 Sequence
15	c	13.6	68.0	45	6	AR200265	AR200265 Sequence
16	c	13.6	68.0	45	6	I27789	I27789 Sequence 21
17	c	13.6	68.0	45	6	I27790	I27790 Sequence 22
18	c	13.6	68.0	80	9	HSD5S347	X64597 H.sapiens D
19		13.6	68.0	87	6	AX282658	AX282658 Sequence
20	c	13.4	67.0	32	6	AR083991	AR083991 Sequence
21	c	13.2	66.0	27	6	AR063107	AR063107 Sequence
22	c	13.2	66.0	27	6	AR122634	AR122634 Sequence
23	c	13.2	66.0	27	6	AR164262	AR164262 Sequence
24	c	13.2	66.0	27	6	AR208934	AR208934 Sequence
25	c	13.2	66.0	28	6	AX397893	AX397893 Sequence
26	c	13.2	66.0	29	6	I56811	I56811 Sequence 1
27	c	13.2	66.0	78	9	HUMRGMB	K03434 Human 28S r
28	c	13.2	66.0	78	10	S69304S2	S69224 p53 (intron
29	c	13.2	66.0	89	6	AX010661	AX010661 Sequence
30		13.2	66.0	89	6	AX010662	AX010662 Sequence
31		13.2	66.0	92	5	AF173802	AF173802 Chloephag
32	c	13.2	66.0	92	11	AF042209	AF042209 Bos tauru
33	c	13	65.0	31	6	BD002386	BD002386 Gene comp
34	c	12.8	64.0	20	6	AX298446	AX298446 Sequence
35		12.8	64.0	51	6	AR117105	AR117105 Sequence
36		12.8	64.0	51	6	AX161745	AX161745 Sequence
37		12.8	64.0	80	9	S57152	S57152 Homo sapien
38		12.8	64.0	90	6	AX455995	AX455995 Sequence
39	c	12.6	63.0	22	6	E28452	E28452 Novel pepti
40	c	12.6	63.0	31	6	AR070106	AR070106 Sequence
41		12.6	63.0	36	6	AX393601	AX393601 Sequence
42	c	12.6	63.0	38	6	A23247	A23247 oligonucleo
43	c	12.6	63.0	38	6	A23248	A23248 oligonucleo
44	c	12.6	63.0	38	6	AR040879	AR040879 Sequence
45	c	12.6	63.0	38	6	AR040880	AR040880 Sequence

ALIGNMENTS

RESULT 1
AR035484
LOCUS AR035484 60 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 56 from patent US 5871902.
ACCESSION AR035484
VERSION AR035484.1 GI:5952152
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 60)
AUTHORS Weininger,S. and Weininger,A.M.
TITLE Sequence-specific detection of nucleic acid hybrids using a DNA-binding molecule or assembly capable of discriminating perfect hybrids from non-perfect hybrids

```
  JOURNAL   Patent: US 5871902-A 56 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="unknown"
BASE COUNT       16 a      23 c      13 g       8 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 60;
  Best Local Similarity   84.2%;  Pred. No. 2.8e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACCCCACAGACAGCCCCC 20
           ||| ||||||||| | |||
Db      15 GACACCACAGACACCGCCC 33


RESULT 2
AR003361/c
LOCUS       AR003361                  25 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 8 from patent US 5744303.
ACCESSION   AR003361
VERSION     AR003361.1  GI:3964620
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Iggo,R., Friend,S.H., Frebourg,T. and Ishioka,C.
  TITLE     Functional assay for transcriptional regulator genes
  JOURNAL   Patent: US 5744303-A 8 28-APR-1998;
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="unknown"
BASE COUNT        6 a       3 c      12 g       4 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 25;
  Best Local Similarity   88.2%;  Pred. No. 4.9e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCCCACAGACAGCCCCC 20
           |||||| ||||||| ||
Db      20 CCCCACTGACAGCCTCC 4


RESULT 3
AR168794/c
LOCUS       AR168794                  26 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 20 from patent US 6288042.
ACCESSION   AR168794
VERSION     AR168794.1  GI:17904898
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Rando,R.F., Ojwaug,J.O., Hogan,M.E., Wallace,T.L. and Cossum,P.A.
  TITLE     Anti-viral guanosine-rich tetrad forming oligonucleotides
  JOURNAL   Patent: US 6288042-A 20 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        0 a       0 c      17 g       9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           | |||||||  ||| |||||
Db      23 CCACCCCACCAACAACCCCC 4


RESULT 4
AR168799/c
LOCUS       AR168799                  26 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 25 from patent US 6288042.
ACCESSION   AR168799
VERSION     AR168799.1  GI:17904906
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Rando,R.F., Ojwaug,J.O., Hogan,M.E., Wallace,T.L. and Cossum,P.A.
  TITLE     Anti-viral guanosine-rich tetrad forming oligonucleotides
  JOURNAL   Patent: US 6288042-A 25 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        0 a       0 c      17 g       9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           | |||||||  ||| |||||
Db      23 CCACCCCACCAACAACCCCC 4


RESULT 5
AR168801/c
LOCUS       AR168801                  26 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 27 from patent US 6288042.
ACCESSION   AR168801
VERSION     AR168801.1  GI:17904909
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Rando,R.F., Ojwaug,J.O., Hogan,M.E., Wallace,T.L. and Cossum,P.A.
  TITLE     Anti-viral guanosine-rich tetrad forming oligonucleotides
  JOURNAL   Patent: US 6288042-A 27 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        0 a       0 c      17 g       9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           | |||||||  ||| |||||
Db      23 CCACCCCACCAACAACCCCC 4


RESULT 6
AR200263/c
LOCUS       AR200263                  26 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 20 from patent US 6355785.
ACCESSION   AR200263
VERSION     AR200263.1  GI:20250337
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Rando,R.F., Fennewald,S., Zendegui,J.G., Ojwang,J.O., Hogan,M.E.,
            Pommier,Y. and Mazumder,A.
  TITLE     Guanosine-rich oligonucleotide integrase inhibitors
```

```
  JOURNAL   Patent: US 6355785-A 20 12-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        0 a      0 c     17 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 7
AR200268/c
LOCUS       AR200268                 26 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 25 from patent US 6355785.
ACCESSION   AR200268
VERSION     AR200268.1  GI:20250342
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Rando,R.F., Fennewald,S., Zendegui,J.G., Ojwang,J.O., Hogan,M.E.,
            Pommier,Y. and Mazumder,A.
  TITLE     Guanosine-rich oligonucleotide integrase inhibitors
  JOURNAL   Patent: US 6355785-A 25 12-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        0 a      0 c     17 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 8
AR200270/c
LOCUS       AR200270                 26 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 27 from patent US 6355785.
ACCESSION   AR200270
VERSION     AR200270.1  GI:20250344
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Rando,R.F., Fennewald,S., Zendegui,J.G., Ojwang,J.O., Hogan,M.E.,
            Pommier,Y. and Mazumder,A.
  TITLE     Guanosine-rich oligonucleotide integrase inhibitors
  JOURNAL   Patent: US 6355785-A 27 12-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        0 a      0 c     17 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 9
I27788/c
LOCUS       I27788                   26 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 20 from patent US 5567604.
ACCESSION   I27788
VERSION     I27788.1  GI:1818564
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Rando,R.F., Fennewald,S., Zendegui,J.G. and Ojwang,J.O.
  TITLE     Anti-viral guanosine-rich oligonucleotides
  JOURNAL   Patent: US 5567604-A 20 22-OCT-1996;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        0 a      0 c     17 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 10
I27793/c
LOCUS       I27793                   26 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 25 from patent US 5567604.
ACCESSION   I27793
VERSION     I27793.1  GI:1818569
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Rando,R.F., Fennewald,S., Zendegui,J.G. and Ojwang,J.O.
  TITLE     Anti-viral guanosine-rich oligonucleotides
  JOURNAL   Patent: US 5567604-A 25 22-OCT-1996;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        0 a      0 c     17 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 11
I27795/c
LOCUS       I27795                   26 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 27 from patent US 5567604.
ACCESSION   I27795
VERSION     I27795.1  GI:1818571
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Rando,R.F., Fennewald,S., Zendegui,J.G. and Ojwang,J.O.
```

```
  TITLE     Anti-viral guanosine-rich oligonucleotides
  JOURNAL   Patent: US 5567604-A 27 22-OCT-1996;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        0 a      0 c     17 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     23 CCACCCCACCAACAACCCCC 4


RESULT 12
AR168795/c
LOCUS       AR168795                  45 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 21 from patent US 6288042.
ACCESSION   AR168795
VERSION     AR168795.1  GI:17904900
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Rando,R.F., Ojwaug,J.O., Hogan,M.E., Wallace,T.L. and Cossum,P.A.
  TITLE     Anti-viral guanosine-rich tetrad forming oligonucleotides
  JOURNAL   Patent: US 6288042-A 21 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT        0 a      0 c     30 g     15 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 45;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     42 CCACCCCACCAACAACCCCC 23


RESULT 13
AR168796/c
LOCUS       AR168796                  45 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 22 from patent US 6288042.
ACCESSION   AR168796
VERSION     AR168796.1  GI:17904901
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Rando,R.F., Ojwaug,J.O., Hogan,M.E., Wallace,T.L. and Cossum,P.A.
  TITLE     Anti-viral guanosine-rich tetrad forming oligonucleotides
  JOURNAL   Patent: US 6288042-A 22 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT        0 a      0 c     30 g     15 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 45;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     42 CCACCCCACCAACAACCCCC 23


RESULT 14
AR200264/c
LOCUS       AR200264                  45 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 21 from patent US 6355785.
ACCESSION   AR200264
VERSION     AR200264.1  GI:20250338
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Rando,R.F., Fennewald,S., Zendegui,J.G., Ojwang,J.O., Hogan,M.E.,
            Pommier,Y. and Mazumder,A.
  TITLE     Guanosine-rich oligonucleotide integrase inhibitors
  JOURNAL   Patent: US 6355785-A 21 12-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT        0 a      0 c     30 g     15 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 45;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     42 CCACCCCACCAACAACCCCC 23


RESULT 15
AR200265/c
LOCUS       AR200265                  45 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 22 from patent US 6355785.
ACCESSION   AR200265
VERSION     AR200265.1  GI:20250339
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Rando,R.F., Fennewald,S., Zendegui,J.G., Ojwang,J.O., Hogan,M.E.,
            Pommier,Y. and Mazumder,A.
  TITLE     Guanosine-rich oligonucleotide integrase inhibitors
  JOURNAL   Patent: US 6355785-A 22 12-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unknown"
BASE COUNT        0 a      0 c     30 g     15 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 45;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGACCCCACAGACAGCCCCC 20
          | ||||||| ||| |||||
Db     42 CCACCCCACCAACAACCCCC 23


Search completed: December  3, 2002, 18:13:17
Job time : 308.2 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-12
Perfect score: 20
Sequence: 1 cgaccccacagacagccccc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	21	AAZ31442	Human neuropilin m
2		14.4	72.0	60	24	ABN46092	Human spliced tran
3	c	14.2	71.0	37	24	ABK81254	Polyimmunoglobulin
4	c	14.2	71.0	49	13	AAQ33751	Upstream sequence
5		14.2	71.0	60	17	AAT30628	Probe nucleic acid
6		14	70.0	45	22	AAC82189	Human retrovirus D
7	c	13.8	69.0	25	15	AAQ61824	Primer for mutant
8		13.8	69.0	30	21	AAA40611	SHR sequence varia
9		13.8	69.0	60	21	AAA05788	Streptavidin displ
10		13.6	68.0	22	21	AAA53503	NF-kappaB binding
11	c	13.6	68.0	26	15	AAQ79225	Guanosine rich oli
12	c	13.6	68.0	26	15	AAQ79227	Guanosine rich oli
13	c	13.6	68.0	26	15	AAQ79220	Guanosine rich oli
14	c	13.6	68.0	26	18	AAT51645	Viral integrase in
15	c	13.6	68.0	26	18	AAT51638	Viral integrase in
16	c	13.6	68.0	26	18	AAT51643	Viral integrase in
17	c	13.6	68.0	26	19	AAX79227	Oligonucleotide #2
18	c	13.6	68.0	45	15	AAQ79221	Guanosine rich oli
19	c	13.6	68.0	45	15	AAQ79222	Guanosine rich oli
20	c	13.6	68.0	45	18	AAT51639	Viral integrase in
21	c	13.6	68.0	45	18	AAT51640	Viral integrase in
22	c	13.6	68.0	45	19	AAX79228	Oligonucleotide #2
23		13.6	68.0	45	21	AAA05784	Streptavidin displ
24		13.6	68.0	65	24	ABN56220	Mouse spliced tran
25		13.6	68.0	87	21	AAC10414	Human secreted pro
26		13.6	68.0	87	22	AAH43509	C1435G SNP from DN
27	c	13.4	67.0	22	21	AAA96906	RACE PCR primer us
28		13.4	67.0	40	20	AAV82040	Moraxella lactofer
29	c	13.4	67.0	60	24	ABN32297	Human spliced tran
30	c	13.4	67.0	89	18	AAT76219	Human IL5 antisens
31	c	13.4	67.0	89	20	AAX54024	Human IL-5 antisen
32	c	13.4	67.0	89	21	AAF19590	Human IL5 polynucl
33	c	13.4	67.0	89	21	AAA33468	Low adenosine anti
34		13.2	66.0	18	24	AAS18236	Rat Arginase I PCR
35	c	13.2	66.0	27	22	AAH76134	Human Shh CA repea
36	c	13.2	66.0	27	22	AAD10173	Human Sonic hedgeh
37	c	13.2	66.0	27	22	AAC87099	PCR primer for cDN
38	c	13.2	66.0	27	24	ABN87571	Human sonic hedgeh
39	c	13.2	66.0	28	24	ABK52396	SahH gene encoding
40	c	13.2	66.0	29	17	AAT07256	Oligonucleotide in
41	c	13.2	66.0	29	17	AAT08788	Cytokine productio
42	c	13.2	66.0	32	24	ABK70089	GMD PCR primer #7
43	c	13.2	66.0	60	24	ABN35987	Human spliced tran
44		13.2	66.0	60	24	ABN43371	Human spliced tran
45		13.2	66.0	90	22	ABA48830	Human breast cell

ALIGNMENTS

RESULT 1
AAZ31442
ID AAZ31442 standard; DNA; 20 BP.
XX
AC AAZ31442;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3612.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 16; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 5 A; 12 C; 3 G; 0 U; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGACCCCACAGACAGCCCCC 20
||||||||||||||||||||
Db 1 CGACCCCACAGACAGCCCCC 20

RESULT 2
ABN46092
ID ABN46092 standard; DNA; 60 BP.
XX
AC ABN46092;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:18840.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 18840; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 15 A; 17 C; 14 G; 14 T; 0 other;

Query Match 72.0%; Score 14.4; DB 24; Length 60;
Best Local Similarity 93.8%; Pred. No. 2.6e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGACCCCACAGACAGC 16
||||||||||||| ||
Db 31 CGACCCCACAGACTGC 46

RESULT 3
ABK81254/c
ID ABK81254 standard; DNA; 37 BP.
XX
AC ABK81254;
XX
DT 13-AUG-2002 (first entry)
XX
DE Polyimmunoglobulin receptor (pIgR) related primer #44.
XX
KW Transcellular transport; transcytotic transport; paracellular transport;
KW respiratory system disorder; lung cancer; tumour; asthma;
KW pathogenic infection; allergy-related disorder;
KW gastrointestinal tract disorder; gastrointestinal hormone disorder;
KW Chron's disease; eating disorder; polyimmunoglobulin receptor; pIgR;
KW primer; ss.
XX
OS Synthetic.
XX
PN WO200228408-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30832.
XX
PR 02-OCT-2000; 2000US-237929P.
PR 13-NOV-2000; 2000US-248478P.
PR 14-NOV-2000; 2000US-248819P.
PR 09-FEB-2001; 2001US-267601P.
XX
PA (ARIZ-) ARIZEKE PHARM INC.
XX
PI Houston LL, Sheridan PJ, Hawley S, Glynn JM, Chapin S, Basu A;
XX
DR WPI; 2002-416628/44.
XX
PT Complex useful for transporting active agent through epithelial
PT barrier, has biologically active portion and target element directed to
PT ligand that confers e.g. transcytotic properties to agent specific to
PT ligand -
XX
PS Example 3; Fig 12B; 379pp; English.
XX
CC The invention described a complex or compound (I) comprising a
CC biologically active portion and a target element (II) directed to a

CC ligand that confers transcellular, transcytotic or paracellular
CC transporting properties to an agent specifically bound to the ligand,
CC where (II) is not an antibody. Alternatively, (I) comprises two or more
CC (II) directed to one or more ligands. (I) is useful for delivering a
CC biologically active agent to an animal, for transporting an active agent
CC through an epithelial or mucosal barrier, and for treating or
CC identifying a disease in an animal e.g. diseases of the respiratory
CC system including lung cancer and tumours, asthma, pathogenic infections,
CC allergy-related disorders, gastrointestinal tract disorders, disorders
CC relating to gastrointestinal hormones, Chron's disease, eating disorders
CC and any disease or disorder involving polyimmunoglobulin receptor (pIgR)
CC displaying cells. This sequence represents a primer associated with the
CC transport of biologically active agents across cellular barriers.
XX
SQ Sequence 37 BP; 3 A; 13 C; 14 G; 7 T; 0 other;

Query Match 71.0%; Score 14.2; DB 24; Length 37;
Best Local Similarity 84.2%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GACCCCACAGACAGCCCCC 20
||| ||||||| ||||||
Db 28 GACGCCACAGAACGCCCCC 10

RESULT 4
AAQ33751/c
ID AAQ33751 standard; DNA; 49 BP.
XX
AC AAQ33751;
XX
DT 02-FEB-1993 (first entry)
XX
DE Upstream sequence of microsatellite from clone TGLA162.
XX
KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
KW genetic mapping; traits; amplification; ss.
XX
OS Bos taurus.
XX
PN WO9213102-A.
XX
PD 06-AUG-1992.
XX
PF 15-JAN-1992; 92WO-US00340.
XX
PR 15-JAN-1991; 91US-0642342.
XX
PA (GENM-) GENMARK.
XX
PI Georges M, Massey JM;
XX
DR WPI; 1992-284684/34.
XX
PT Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding
XX
PS Table 7; Page 229; 517pp; English.
XX
CC The sequence is that upstream of a bovine microsatellite sequence
CC obtd. by screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (T6)n >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPTIPRIM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also AAQ33501-34437.
XX
SQ Sequence 49 BP; 12 A; 11 C; 16 G; 10 T; 0 other;

Query Match 71.0%; Score 14.2; DB 13; Length 49;
Best Local Similarity 84.2%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGACCCCACAGACAGCCCC 19
|||||||| |||| || ||
Db 19 CGACCCCATAGACGGCACC 1

RESULT 5
AAT30628
ID AAT30628 standard; cDNA; 60 BP.
XX
AC AAT30628;
XX
DT 21-FEB-1997 (first entry)
XX
DE Probe nucleic acid #11.
XX
KW Target binding assembly; nucleic acid recognition unit; NF-kappa-B; SP1;
KW TATA; human papillomavirus; HPV; HPV E2; human immunodeficiency virus;
KW HIV LTR; Tat binding unit; probe; assembly sequence; asymmetry sequence;
KW nuclear localisation signal sequence; human cell; HPV LTR; HIV; bacteria;
KW virus; ss.
XX
OS Synthetic.
XX
PN WO9617956-A2.
XX
PD 13-JUN-1996.
XX
PF 07-DEC-1995; 95WO-US15944.
XX
PR 09-DEC-1994; 94US-0353476.
XX
PA (GENE-) GENE POOL INC.
XX
PI Weininger AM, Weininger S;
XX
DR WPI; 1996-287199/29.
XX
PT Probe nucleic acids, target binding assemblies, etc - for detection
PT and localisation of specific nucleic acid sequences, esp. HIV and
PT HPV
XX
PS Disclosure; Page 82; 172pp; English.
XX
CC AAT30615-T30634 represent probe nucleic acids of the invention. The
CC probe of the invention contains a target binding region (TBR), a booster
CC binding region (BBR), and an optional support or attachment (OSA). The
CC target binding assembly (TBA) recognised by the TBR (see
CC AAT30581-T30614) of the probe, contains at least one nucleic acid
CC recognition unit (NAR), and optionally a linker sequence, an assembly
CC sequence (see AAR95994-R95998), an asymmetry sequence (see
CC AAR95999-R96006), a nuclear localisation signal sequence (see AAR96007),
CC and an OSA. The assembly sequence and asymmetry sequences are
CC responsible for the folding and association of the NARs. The NARs (see
CC AAR95965-R95993) are selected from NF-kappa-B, SP1, TATA, human
CC papillomavirus (HPV) E2, HPV LTR, human immunodeficiency virus (HIV) LTR
CC and Tat binding units. The linker sequence is an oligopeptide, which
CC does not interfere with NAR function, but provides stability and control
CC over the spacing of the NAR from the rest of the TBA. The OSA is an
CC attached support or indicator, or other means of localisation of the
CC probe. The probe can be used in a method for detecting or localising a
CC specific target nucleic acid sequence (TNA). The method is highly
CC sensitive, and has a high degree of specificity. The method can be used
CC for detecting specific nucleic acid sequences, including those found in
CC human cells, in HIV, HPV, and other nucleic acid containing systems,

```
CC   including bacteria and viruses.
XX
SQ   Sequence 60 BP; 16 A; 23 C; 13 G; 8 T; 0 other;

  Query Match            71.0%;  Score 14.2;  DB 17;  Length 60;
  Best Local Similarity  84.2%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACCCCACAGACAGCCCCC 20
           ||| ||||||||| | |||
Db      15 GACACCACAGACACCGCCC 33


RESULT 6
AAC82189
ID   AAC82189 standard; DNA; 45 BP.
XX
AC   AAC82189;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Human retrovirus DNA probe 2J #1.
XX
KW   Detection; infection; disease; diagnosis; risk assessment; probe;
KW   PCR primer; ss.
XX
OS   Human retrovirus.
XX
PN   DE19921419-A1.
XX
PD   16-NOV-2000.
XX
PF   08-MAY-1999;   99DE-1021419.
XX
PR   08-MAY-1999;   99DE-1021419.
XX
PA   (UYHE-) UNIV HEIDELBERG RUPRECHT-KARLS.
XX
PI   Seifarth W,  Leib-Moesch C,  Baust C;
XX
DR   WPI; 2001-000507/01.
XX
PT   Specific detection of retroviral nucleic acid, useful e.g. for early
PT   diagnosis of disease, comprises amplification with universal primers
PT   then reverse dot blot hybridization  -
XX
PS   Example 3; Page 11; 22pp; German.
XX
CC   This invention describes a novel method for the specific detection and
CC   identification of retroviral nucleic acid (or retroviruses) comprising
CC   isolation of DNA and/or RNA, subjecting this to reverse
CC   transcription-polymerase chain reaction (RT-PCR) with one of two
CC   specified primer pairs (or both pairs), purifying the amplicons and
CC   analyzing them by reverse dot blot hybridization (RDBH) using
CC   immobilized, synthetic oligonucleotide probes. The method is used to
CC   detect retroviruses (or their nucleic acid) in cell cultures (or
CC   supernatants); body samples or other biological materials. It produces an
CC   expression pattern of all retroviruses in a single experiment,
CC   particularly for large scale surveys to determine any correlations
CC   between diseases and activity of selected retroviruses. If such a
CC   correlation is found, the method will allow early diagnosis, or
CC   assessment of risk. An efficient, reliable and rapid method for detecting
CC   all presently known endogenous and exogenous retroviruses of human or
CC   animal origin with very low detection limit. The amplification primers
CC   are universal and inclusion of a clamp and restriction enzyme recognition
CC   site in them (i) improves primer/matrix binding kinetics and allows
CC   amplification of RTase templates even if there is not an exact match and
CC   (ii) facilitates subsequent cloning. Using probes that do not overlap
CC   with primers overcomes the problem of amplicons reacting, to some extent,
CC   with all probes in the array.
XX
SQ   Sequence 45 BP; 12 A; 17 C; 4 G; 12 T; 0 other;

  Query Match            70.0%;  Score 14;  DB 22;  Length 45;
  Best Local Similarity  100.0%;  Pred. No. 3.9e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 CACAGACAGCCCCC 20
           ||||||||||||||
Db       3 CACAGACAGCCCCC 16


RESULT 7
AAQ61824/c
ID   AAQ61824 standard; DNA; 25 BP.
XX
AC   AAQ61824;
XX
DT   19-SEP-1994  (first entry)
XX
DE   Primer for mutant p53 sequence.
XX
KW   Fragment A; RGC; transcriptional regulator; TR; germ line mutation;
KW   assay; cancer; PCR; amplification; sense; ss.
XX
OS   Synthetic.
XX
PN   WO9408049-A.
XX
PD   14-APR-1994.
XX
PF   28-SEP-1993;   93WO-US09259.
XX
PR   01-OCT-1992;   92US-0956696.
PR   12-APR-1993;   93US-0046033.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (SURE-) INST SUISSE RECH EXPERIMENTALE.
XX
PI   Frebourg T,  Friend SH,  Iggo R,  Ishioka C;
XX
DR   WPI; 1994-135609/16.
XX
PT   Assay for mutations in a transcriptional regulator gene - using
PT   cells transfected with the gene and DNA encoding a detectable
PT   protein expressed in response to the gene.
XX
PS   Example; Page 21; 51pp; English.
XX
CC   The sequence is that of a sense PCR primer for amplification of a
CC   mutant p53 gene.  The PCR prod. can be used in an assay to directly
CC   assess whether a germ-line mutation in a gene which is a trans-
CC   criptional regulator results in the prodn. of a non functional gene.
CC   This may be used in screening individuals at risk of cancer or in cancer
CC   prognosis.
CC   See also AAQ61818-23.
XX
SQ   Sequence 25 BP; 6 A; 3 C; 12 G; 4 T; 0 other;

  Query Match            69.0%;  Score 13.8;  DB 15;  Length 25;
  Best Local Similarity  88.2%;  Pred. No. 4.7e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCCCACAGACAGCCCCC 20
           |||||| ||||||| ||
Db      20 CCCCACTGACAGCCTCC 4


RESULT 8
AAA40611
ID   AAA40611 standard; DNA; 30 BP.
XX
AC   AAA40611;
XX
DT   15-AUG-2000  (first entry)
XX
```

```
DE   SHR sequence variant number 2 SEQ ID NO:11.
XX
KW   Human; rat; CD36; SHR; spontaneous hypertensive rat; diagnosis;
KW   therapy; screening; polymorphism; variant; detection; mutant;
KW   blood; mutation; insulin; glucose metabolism; fatty acid metabolism;
KW   catecholamine; malaria; infection; parasite; antiparasitic;
KW   antidiabetic; primer; ss.
XX
OS   Rattus sp.
OS   Synthetic.
XX
PN   WO200019883-A2.
XX
PD   13-APR-2000.
XX
PF   07-OCT-1999;   99WO-US23418.
XX
PR   07-OCT-1998;   98US-0167750.
PR   28-DEC-1998;   98US-0221222.
PR   17-MAR-1999;   99US-0270542.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
PA   (SCIO-) SCIOS INC.
PA   (AITM/) AITMAN T J.
PA   (SCOT/) SCOTT J.
PA   (STAN/) STANTON L W.
XX
PI   Aitman TJ,  Scott J,  Stanton LW;
XX
DR   WPI; 2000-303596/26.
XX
PT   Nucleic acids encoding mutant CD36 proteins useful for preventing,
PT   diagnosing and treating parasitic infections, especially malaria -
XX
PS   Example 4; Page 83; 167pp; English.
XX
CC   The present invention describes isolated nucleic acid molecules (A)
CC   encoding mutant CD36 proteins (B). Parasites such as Plasmodium
CC   falciparum (the major cause of malaria) are unable to utilise the
CC   mutated proteins to gain entry to, and infect cells. The mutant CD36
CC   proteins do not function correctly preventing parasites utilising them
CC   to infect cells. The nucleic acids may be used for the recombinant
CC   production of mutant CD36 proteins according to standard methodologies.
CC   They may be used in this way to prevent and treat parasitic infections
CC   that utilise the CD36 protein to infect cells, such as P. falciparum,
CC   the major cause of malaria. For example, the protein may be used to
CC   identify modulators of CD36 expression and activity or a patient's CD36
CC   DNA may be screened to determine whether there are any mutations present
CC   that may confer resistance to parasitic infections. The proteins and
CC   nucleic acids may also be used to prevent, diagnose and treat diseases
CC   associated with defects in insulin action and/or glucose metabolism
CC   and/or fatty acid metabolism and/or catecholamine action in subjects
CC   possessing mutations in the CD36 genes. AAA40606 to AAA40759, and
CC   AAB02515 to AAB02564, represent nucleotide and amino acid sequences
CC   respectively which are used in the exemplification of the present
CC   invention.
XX
SQ   Sequence 30 BP; 10 A; 10 C; 5 G; 5 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 21;  Length 30;
  Best Local Similarity   88.2%;  Pred. No. 4.7e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GACCCCACAGACAGCCC 18
           ||||||| ||||||| |
Db       9 GACCCCAAAGACAGCAC 25


RESULT 9
AAA05788
ID   AAA05788 standard; DNA; 60 BP.
XX
AC   AAA05788;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Streptavidin display linker DNA, SEQ ID NO:250.
XX
KW   Phage display; bacteriophage M13; fusion protein; major coat protein;
KW   protein VIII; phagemid vector; electroporation; combinatorial library;
KW   streptavidin; SAV; ss.
XX
OS   Synthetic.
XX
PN   WO200006717-A2.
XX
PD   10-FEB-2000.
XX
PF   22-JUL-1999;   99WO-US16596.
XX
PR   27-JUL-1998;   98US-0094291.
PR   08-OCT-1998;   98US-0103514.
PR   10-MAY-1999;   99US-0133296.
PR   19-MAY-1999;   99US-0134870.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Sidhu SS,  Weiss GA,  Wells JA;
XX
DR   WPI; 2000-183122/16.
DR   P-PSDB; AAY81345.
XX
PT   Fusion proteins comprising a heterologous protein and a viral variant
PT   major coat protein useful in phage display systems for improving
PT   transformation efficiency  -
XX
PS   Disclosure; Fig 7B; 118pp; English.
XX
CC   The invention relates to novel fusion proteins comprising a heterologous
CC   polypeptide fused to a variant (non-wild type) bacteriophage
CC   major coat protein (protein VIII). The invention also relates to
CC   replicable expression vectors which contain a gene encoding the fusion
CC   protein; host cells containing the expression vectors; phages which
CC   display the fusion protein on their surface; phage libraries displaying
CC   a plurality of different fusion proteins on viral surfaces; and methods
CC   of using these compositions. The fusion proteins the invention are well
CC   tolerated in phage display systems. Variants of the major coat proteins
CC   can be used to alter the number of fusion proteins incorporated into a
CC   virus particle. Hyper-functional variants can be used to increase the
CC   number of fusion proteins incorporated into a virus particle.
CC   Conversely, hypo-functional variants can be used to decrease fusion
CC   protein incorporation. This is useful for tailoring the incorporation of
CC   fusion proteins into virus particles to achieve a desired level of
CC   valency. The variant replicable plasmid/phagemid vectors are useful for
CC   producing polypeptides of interest. The methods are useful for improving
CC   the transformation of cells by highly purifying DNA. The present
CC   invention uses affinity DNA purification to reduce ionic impurities and
CC   thus reduce the conductance associated with a unit mass of DNA. This is
CC   an advantageous in electroporation methods for increasing the
CC   concentration of DNA present. The increase in DNA entering the host cell
CC   provides a greater number of transformants per electroporation and allows
CC   one to prepare larger combinatorial libraries which overcomes the prior
CC   art problem of small library size using recombinant DNA. Sequences
CC   AAA05768-A05795 represent DNA encoding linkers selected for display
CC   of streptavidin (SAV).
XX
SQ   Sequence 60 BP; 10 A; 44 C; 6 G; 0 U; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 21;  Length 60;
  Best Local Similarity   88.2%;  Pred. No. 4.8e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCCCACAGACAGCCCCC 20
           |||||  ||||||||||
Db       6 CCCCAGCGACAGCCCCC 22
```

```
RESULT 10
AAA53503
ID   AAA53503 standard; DNA; 22 BP.
XX
AC   AAA53503;
XX
DT   05-OCT-2000  (first entry)
XX
DE   NF-kappaB binding phosphorotetrathioate aptamer #3.
XX
KW   Phosphorotetrathioate aptamer; NF-kappaB; bacterial pathogenesis; AIDS;
KW   SV-40; rheumatoid arthritis; Crohn's disease; asbestos lung disease;
KW   generalised inflammatory bowel disease; Hodgkin's disease; cancer;
KW   lung injury; atherosclerosis; cytomegalovirus; herpes simplex virus;
KW   rhinovirus; influenza; neurological disorder; JCV; ds.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base    1..22
FT                    /*tag=  a
FT                    /note= "all A nucleotides are 5' monophosphorothioated"
XX
PN   WO200024404-A1.
XX
PD   04-MAY-2000.
XX
PF   26-OCT-1999;   99WO-US24058.
XX
PR   26-OCT-1998;   98US-0105600.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Gorenstein DG,  Aronson J,  Luxon B,  Herzog N;
XX
DR   WPI; 2000-350578/30.
XX
PT   Preparing achiral nuclear factor-kappaB specific aptamers for
PT   generating high binding, nuclease resistant aptamers that retain their
PT   specificity comprises synthesizing amplifying random phosphodiester
PT   oligonucleotide combinatorial library -
XX
PS   Example 3; Page 31; 63pp; English.
XX
CC   The present sequence is a randomly generated sequence created so that
CC   each A residue is 5' monophosphorothioated. It has a high affinity for
CC   the p65 homodimer, which forms part of the nuclear factor kappaB
CC   (NF-kappaB) transcription factor. This sequence can be used in the
CC   production of aptamers which are stable and nuclease resistant, and can
CC   be used to treat diseases involving the transcription factor, such as
CC   bacterial pathogenesis, rheumatoid arthritis, Crohn's disease,
CC   generalised inflammatory bowel disease, asbestos lung disease, Hodgkin's
CC   disease, prostate cancer, ventilator induced lung injury, cancer, AIDS,
CC   human cutaneous T-cell lymphoma, lymphoid malignancies, HTLV-1 induced
CC   adult T-cell leukaemia, atherosclerosis, cytomegalovirus, herpes simplex
CC   virus, JCV, SV-40, rhinovirus, influenza, neurological disorders and
CC   lymphomas.
XX
SQ   Sequence 22 BP; 4 A; 16 C; 2 G; 0 U; 0 other;

  Query Match         68.0%;  Score 13.6;  DB 21;  Length 22;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           || || |||| || ||||||
Db       3 CGCCCACACACACCGCCCCC 22


RESULT 11
AAQ79225/c
ID   AAQ79225 standard; DNA; 26 BP.
XX
AC   AAQ79225;
XX
DT   17-JUL-1995  (first entry)
XX
DE   Guanosine rich oligonucleotide used to treat viral infection.
XX
KW   Guanosine; tetrad; inhibition; replication; virus; treatment;
KW   therapy; infection; herpes simplex virus; human papilloma virus;
KW   Epstein-Barr virus; HIV, adenovirus; respiratory syncitial virus;
KW   hepatitis B virus; human cytomegalovirus; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    26
FT                   /*tag=  a
FT                   /mod_base=
FT                   /note= "Cholesterol moiety attached to this base."
XX
PN   WO9425037-A.
XX
PD   10-NOV-1994.
XX
PF   25-APR-1994;   94WO-US04529.
XX
PR   23-APR-1993;   93US-0053027.
PR   28-OCT-1993;   93US-0145704.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
PA   (TRIP-) TRIPLEX PHARM CORP.
XX
PI   Fennewald S,  Hogan ME,  Ojwang JO,  Rando RF,  Zendegui JG;
XX
DR   WPI; 1994-357890/44.
XX
PT   Oligo-nucleotide(s) rich in guanosine which form guanosine
PT   tetrads  - used to treat viral infections, e.g. herpes-virus and
PT   HIV
XX
PS   Claim 41; Page 56; 101pp; English.
XX
CC   The oligonucleotides (See AAQ79201-52)  can be used to treat viral
CC   infections. The oligonucleotides inhibit viral replication by
CC   forming guanosine tetrads which form a stabilised 3D structure.
CC   Preferred oligonucleotides contain at least 2 runs of at least 2
CC   guanosine bases and may be capped at the 3' terminus with a modifier
CC   selected from polyamine, poly-L-lysine, cholesterol and
CC   propanolamine.  They may also have a modified phosphodiester linkage
CC   or be modified to contain a phosphorothioate linkage.  They are used
CC   to treat infections with viruses such as herpes simplex virus, human
CC   papilloma virus, Epstein-Barr virus, HIV, adenovirus, respiratory
CC   syncitial virus, hepatitis B virus or human cytomegalovirus.
XX
SQ   Sequence 26 BP; 0 A; 0 C; 17 G; 9 T; 0 other;

  Query Match         68.0%;  Score 13.6;  DB 15;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           | ||||||  ||| |||||
Db      23 CCACCCCACCAACAACCCCC 4


RESULT 12
AAQ79227/c
ID   AAQ79227 standard; DNA; 26 BP.
XX
AC   AAQ79227;
XX
DT   17-JUL-1995  (first entry)
XX
```

```
RESULT 15
US-08-285-936-45/c
; Sequence 45, Application US/08285936
; Patent No. 5728821
;  GENERAL INFORMATION:
;    APPLICANT:  Yelton, Dale
;    APPLICANT:  Glaser, Scott
;    APPLICANT:  Huse, William
;    APPLICANT:  Rosok, Mae J.
;    TITLE OF INVENTION:  No. 5728821el Mutant BR96 Antibodies and
;    TITLE OF INVENTION:  Functional Equivalents Reactive With Human Carcinomas
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  11150 Santa Monica Bldv., Suite 400
;      CITY:  Los Angeles
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  90025-3395
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/285,936
;      FILING DATE:  04-AUG-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adriano, Sarah B.
;      REGISTRATION NUMBER:  34,470
;      REFERENCE/DOCKET NUMBER:  30436.16US01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  310-445-1140
;      TELEFAX:  310-445-9031
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  51 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-285-936-45

  Query Match             66.0%;  Score 13.2;  DB 1;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTG 19
           ||||| | |||||| |||
Db      43 GAGGTTCTGGTGGAGGTG 26


Search completed: November 23, 2002, 06:36:09
Job time : 22.55 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-9
Perfect score: 20
Sequence: 1 tgaggtgcgggtggaagtgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	30	9	US-10-073-353-7	Sequence 7, Appli
	2	14.2	71.0	50	9	US-09-905-291A-151	Sequence 151, App
	3	14.2	71.0	50	10	US-09-909-320-151	Sequence 151, App
	4	14.2	71.0	50	10	US-09-909-088B-151	Sequence 151, App
	5	13.6	68.0	24	10	US-09-820-339A-26	Sequence 26, Appl
c	6	13.2	66.0	25	9	US-09-733-042-4	Sequence 4, Appli
c	7	13.2	66.0	31	10	US-09-801-274-191	Sequence 191, App
	8	13.2	66.0	96	10	US-09-864-761-18926	Sequence 18926, A
	9	13.2	66.0	100	10	US-09-764-877-3114	Sequence 3114, Ap
	10	12.8	64.0	33	10	US-09-873-676-70	Sequence 70, Appl
	11	12.8	64.0	82	10	US-09-923-876-1514	Sequence 1514, Ap
c	12	12.6	63.0	20	10	US-09-755-004-10	Sequence 10, Appl
	13	12.6	63.0	25	10	US-09-866-108-13254	Sequence 13254, A
	14	12.6	63.0	25	10	US-09-866-108-13255	Sequence 13255, A
	15	12.6	63.0	25	10	US-09-866-108-13256	Sequence 13256, A
	16	12.6	63.0	25	10	US-09-866-108-13257	Sequence 13257, A
	17	12.6	63.0	25	10	US-09-866-108-13258	Sequence 13258, A
	18	12.6	63.0	25	10	US-09-866-108-13259	Sequence 13259, A
	19	12.6	63.0	25	10	US-09-866-108-13260	Sequence 13260, A
	20	12.6	63.0	53	10	US-09-911-088-4	Sequence 4, Appli
	21	12.6	63.0	79	10	US-09-815-242-1091	Sequence 1091, Ap
c	22	12.6	63.0	99	10	US-09-827-998-40	Sequence 40, Appl
	23	12.4	62.0	17	10	US-09-934-778-15	Sequence 15, Appl
	24	12.4	62.0	82	10	US-09-864-761-20076	Sequence 20076, A
	25	12.2	61.0	21	12	US-10-016-725-7	Sequence 7, Appli
c	26	12.2	61.0	21	12	US-10-016-725-8	Sequence 8, Appli
	27	12.2	61.0	45	10	US-09-326-447-18	Sequence 18, Appl
	28	12.2	61.0	48	9	US-09-905-291A-370	Sequence 370, App
	29	12.2	61.0	48	10	US-09-909-320-370	Sequence 370, App
	30	12.2	61.0	48	10	US-09-909-088B-370	Sequence 370, App
c	31	12.2	61.0	79	10	US-09-326-447-7	Sequence 7, Appli
	32	12.2	61.0	86	10	US-09-326-447-14	Sequence 14, Appl
	33	12.2	61.0	88	10	US-09-864-761-33204	Sequence 33204, A
c	34	12.2	61.0	92	10	US-09-969-373-1231	Sequence 1231, Ap
	35	12.2	61.0	98	10	US-09-326-447-16	Sequence 16, Appl
c	36	12	60.0	54	9	US-09-991-496-88	Sequence 88, Appl
c	37	12	60.0	54	10	US-09-874-923-88	Sequence 88, Appl
c	38	12	60.0	60	10	US-09-918-063-59	Sequence 59, Appl
	39	12	60.0	60	10	US-09-918-063-61	Sequence 61, Appl
c	40	12	60.0	75	10	US-09-938-700-18	Sequence 18, Appl
c	41	12	60.0	78	10	US-09-938-700-21	Sequence 21, Appl
	42	12	60.0	95	10	US-09-864-761-18519	Sequence 18519, A
c	43	12	60.0	100	10	US-09-983-965-4351	Sequence 4351, Ap
c	44	11.8	59.0	17	10	US-09-969-373-2951	Sequence 2951, Ap
	45	11.8	59.0	19	10	US-09-925-548-51	Sequence 51, Appl

ALIGNMENTS

```
RESULT 1
US-10-073-353-7/c
; Sequence 7, Application US/10073353
; Patent No. US20020168658A1
;  GENERAL INFORMATION:
;       APPLICANT: Sherman M. Weissman
;                  Namadev Baskaran
;       TITLE OF INVENTION: Amplification of Nucleic Acids
;       NUMBER OF SEQUENCES: 8
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Morgan, Lewis & Bockius LLP
;            STREET: 1111 Pennsylvania Ave., NW
;            CITY: Washington
;            STATE: D.C.
;            COUNTRY: USA
;            ZIP: 20004
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: Patentin Release #1.0. Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/073,353
;            FILING DATE: 03-May-2002
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/758,662
;            FILING DATE: 1996-12-07
;            APPLICATION NUMBER: US 08/758,662 (CPA)
;            FILING DATE: 1999-02-17
;            APPLICATION NUMBER: US 09/585,437
;            FILING DATE: 2000-06-02
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Michael S. Tuscan
;            REGISTRATION NUMBER: 43,210
;            REFERENCE/DOCKET NUMBER: 44921-5007-03-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 202-739-3000
;            TELEFAX: 202-739-3001
;  INFORMATION FOR SEQ ID NO: 7:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 30 base pairs
;            TYPE: nucleic acid
```

```
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-073-353-7

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 5.1e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTGC 20
           ||| ||  ||||||||||
Db      25 GAGCTGGTGGTGGAAGTGC 7


RESULT 2
US-09-905-291A-151
; Sequence 151, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/905,291A
;  CURRENT FILING DATE:  2001-07-12
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 151
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: oligonucleotide probe
US-09-905-291A-151

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 50;
  Best Local Similarity   84.2%;  Pred. No. 5.3e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTG 19
           ||| | ||||||||| |||
Db      16 TGAAGGGCGGGTGGAGGTG 34


RESULT 3
US-09-909-320-151
; Sequence 151, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/909,320
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
```

```
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 151
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: oligonucleotide probe
US-09-909-320-151

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 50;
  Best Local Similarity   84.2%;  Pred. No. 5.3e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTG 19
           ||| | ||||||||| |||
Db      16 TGAAGGGCGGGTGGAGGTG 34


RESULT 4
US-09-909-088B-151
; Sequence 151, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/909,088B
;  CURRENT FILING DATE:  2001-07-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 151
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: oligonucleotide probe
US-09-909-088B-151

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 50;
  Best Local Similarity   84.2%;  Pred. No. 5.3e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTG 19
           ||| | ||||||||| |||
Db      16 TGAAGGGCGGGTGGAGGTG 34


RESULT 5
US-09-820-339A-26
; Sequence 26, Application US/09820339A
; Patent No. US20020081652A1
; GENERAL INFORMATION:
;  APPLICANT: FUCHS, Sara
;  APPLICANT:  BARCHAN, Dora
;  APPLICANT:  SOUROUJON, Miriam
;  TITLE OF INVENTION: RECOMBINANT FRAGMENTS OF THE HUMAN ACETYLCHOLINE RECEPTOR AND
;  TITLE OF INVENTION:  FOR TREATMENT OF MYASTHENIA GRAVIS
;  FILE REFERENCE: FUCHS=2A
;  CURRENT APPLICATION NUMBER: US/09/820,339A
;  CURRENT FILING DATE:  1999-11-08
;  PRIOR APPLICATION NUMBER: 09/423,398
```

```
;  PRIOR FILING DATE: 1999-11-08
;  PRIOR APPLICATION NUMBER: PCT/IL98/00211
;  PRIOR FILING DATE: 1998-05-06
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-09-820-339A-26

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 24;
  Best Local Similarity   80.0%;  Pred. No. 9.3e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTGC 20
           |||||| ||||||   ||||
Db       4 TGAGGTCCGGGTGACGGTGC 23


RESULT 6
US-09-733-042-4/c
; Sequence 4, Application US/09733042
; Patent No. US20020168709A1
; GENERAL INFORMATION:
;  APPLICANT:   Hennecke, Frank
;  APPLICANT:  Renner, Wolfgang A.
;  TITLE OF INVENTION:   Replicon Based Activation of Endogenous Genes
;  FILE REFERENCE:   1700.0100001
;  CURRENT APPLICATION NUMBER: US/09/733,042
;  CURRENT FILING DATE:  2000-12-11
;  PRIOR APPLICATION NUMBER:   US 60/169,988
;  PRIOR FILING DATE:   1999-12-10
;  NUMBER OF SEQ ID NOS:   49
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Enh-REV
US-09-733-042-4

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 25;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 AGGTGCGGGTGGAAGTGC 20
           |||||||| |||| || |
Db      22 AGGTGCGGTTGGAGGTAC 5


RESULT 7
US-09-801-274-191/c
; Sequence 191, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 191
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-191

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 31;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGT 18
           || || |||||| |||||
Db      18 TGYGGAGCGGGTCGAAGT 1


RESULT 8
US-09-864-761-18926
; Sequence 18926, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18926
;   LENGTH: 96
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AF060568.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.7
```

```
;   OTHER INFORMATION: NT HIT: AF060568.1, EVALUE 2.00e-47
;   OTHER INFORMATION: EST_HUMAN HIT: AA150052.1, EVALUE 8.20e-01
;   OTHER INFORMATION: SWISSPROT HIT: Q00998, EVALUE 9.20e+00
US-09-864-761-18926

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 96;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 AGGTGCGGGTGGAAGTGC 20
           ||||||  ||||||| ||
Db      41 AGGTGCACGTGGAAGAGC 58


RESULT 9
US-09-764-877-3114
; Sequence 3114, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005
;  CURRENT APPLICATION NUMBER: US/09/764,877
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3114
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-877-3114

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 100;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTG 19
           ||||||| ||| | ||||
Db      12 GAGGTGCAGGTTGCAGTG 29


RESULT 10
US-09-873-676-70
; Sequence 70, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
;  APPLICANT: MacDonald, Nicholas J.
;  APPLICANT:  Sim, Kim L.
;  TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
;  FILE REFERENCE: 05213-0378 (43170-259333)
;  CURRENT APPLICATION NUMBER: US/09/873,676
;  CURRENT FILING DATE:  2001-06-04
;  PRIOR APPLICATION NUMBER: US 60/209,065
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: US 60/289,387
;  PRIOR FILING DATE: 2001-05-08
;  NUMBER OF SEQ ID NOS: 123
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: binding molecule
US-09-873-676-70

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 2.1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GTGCGGGTGGAAGTGC 20
           |||||||||||| || |
Db      17 GTGCGGGTGGAGGTTC 32


RESULT 11
US-09-923-876-1514
; Sequence 1514, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
;  APPLICANT: Lalgudi, Raghunath V.
;  APPLICANT:  Kamigaki, Laura Y. (Ito)
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
;  FILE REFERENCE: PL-0012-1 CON
;  CURRENT APPLICATION NUMBER: US/09/923,876
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: 09/298,329
;  PRIOR FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: 60/085,331
;  PRIOR FILING DATE: 1998-05-05
;  NUMBER OF SEQ ID NOS: 6332
;  SOFTWARE: PERL Program
; SEQ ID NO 1514
;   LENGTH: 82
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020013958A1 700158952H1
;   NAME/KEY: unsure
;   LOCATION: 2, 12-13, 37, 51, 58, 65, 67, 81
;   OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-1514

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 82;
  Best Local Similarity   77.8%;  Pred. No. 2.3e+03;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTG 19
           ||||||   ||||| |||
Db      59 GAGGTGNANGTGGAGGTG 76


RESULT 12
US-09-755-004-10/c
; Sequence 10, Application US/09755004
; Patent No. US20020110810A1
; GENERAL INFORMATION:
;  APPLICANT:   Shuber, Anthony
;  TITLE OF INVENTION:   Methods for Detecting, Grading or Monitoring an H. pylori In
;  FILE REFERENCE:   EXT-048
;  CURRENT APPLICATION NUMBER: US/09/755,004
;  CURRENT FILING DATE:  2001-01-05
;  NUMBER OF SEQ ID NOS:   11
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 10
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: APC forward primer
US-09-755-004-10

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 20;
  Best Local Similarity   78.9%;  Pred. No. 2.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTG 19
           ||||| |  |||||| |||
Db      19 TGAGGAGGTGGTGGAGGTG 1


RESULT 13
US-09-866-108-13254
```

```
; Sequence 13254, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 13254
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-13254

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTGC 20
           |||| || || | ||||||
Db       7 GAGGAGCTGGAGAAAGTGC 25


RESULT 14
US-09-866-108-13255
; Sequence 13255, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 13255
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-13255

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTGC 20
           |||| || || | ||||||
Db       6 GAGGAGCTGGAGAAAGTGC 24


RESULT 15
US-09-866-108-13256
; Sequence 13256, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 13256
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-13256

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTGC 20
           |||| || || | ||||||
Db       5 GAGGAGCTGGAGAAAGTGC 23


Search completed: November 23, 2002, 06:42:07
Job time : 17.25 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-9
Perfect score: 20
Sequence: 1 tgaggtgcgggtggaagtgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.8	74.0	100	9	AI198185	AI198185 qi51f03.x
	2	14.4	72.0	75	17	BH790276	BH790276 SALK_0567
	3	14.2	71.0	61	12	BF507253	BF507253 5923P-29
	4	14.2	71.0	73	10	BE130245	BE130245 L48-385T3
c	5	14.2	71.0	82	9	AA646813	AA646813 vn37b09.r
	6	14.2	71.0	86	12	BG402252	BG402252 602465822
c	7	14.2	71.0	99	9	AI559239	AI559239 tq32e03.x
	8	13.8	69.0	50	9	AU102737	AU102737 AU102737
	9	13.6	68.0	60	17	AZ982649	AZ982649 2M0263H04
	10	13.6	68.0	61	9	AA659447	AA659447 nu24b08.s
	11	13.6	68.0	69	17	DR17P21S	AL740955 Danio rer
	12	13.6	68.0	88	9	AA500773	AA500773 vg01b07.r
c	13	13.6	68.0	92	14	H55398	H55398 CHR220337 C
c	14	13.6	68.0	96	14	Z20250	Z20250 HSAAABRWI P
c	15	13.2	66.0	72	17	AQ248436	AQ248436 T7M22-Sp6
c	16	13.2	66.0	73	12	BG879136	BG879136 ib70h10.y
c	17	13.2	66.0	76	9	AI221944	AI221944 qh01a03.x
	18	13.2	66.0	84	10	AV848064	AV848064 AV848064
	19	13.2	66.0	88	14	W29984	W29984 mc23d02.r1
c	20	13.2	66.0	90	17	BH233613	BH233613 1006174G0
c	21	13.2	66.0	91	17	AF190831	AF190831 AF190831
c	22	13.2	66.0	95	14	BQ753757	BQ753757 EBca01_SQ
	23	13.2	66.0	97	9	AA097217	AA097217 mk10a12.r
c	24	13.2	66.0	99	9	AA704622	AA704622 zf86a04.s
	25	13	65.0	19	17	AZ782026	AZ782026 2M0021I23
	26	13	65.0	31	9	AI021071	AI021071 ua99f05.r
	27	12.8	64.0	50	9	AU104342	AU104342 AU104342
	28	12.8	64.0	61	9	AI877694	AI877694 fc50e10.y
c	29	12.8	64.0	68	13	BI763776	BI763776 603049718
c	30	12.8	64.0	77	10	AW022596	AW022596 df41b06.y
c	31	12.8	64.0	82	14	R69402	R69402 yj83d10.r1
	32	12.8	64.0	96	9	AI263688	AI263688 qx67h11.x
c	33	12.8	64.0	96	17	TA192C04P	AL477388 T. brucei
	34	12.6	63.0	35	17	AZ986475	AZ986475 2M0268N09
c	35	12.6	63.0	43	14	T63536	T63536 yc07c10.s1
c	36	12.6	63.0	47	17	AZ373977	AZ373977 1M0126A05
	37	12.6	63.0	47	17	AZ949602	AZ949602 2M0213H20
c	38	12.6	63.0	48	17	AZ496155	AZ496155 1M0332K15
	39	12.6	63.0	49	14	R53501	R53501 yj70d10.r1
	40	12.6	63.0	50	9	AU106147	AU106147 AU106147
	41	12.6	63.0	57	17	AZ344037	AZ344037 1M0077A19
	42	12.6	63.0	58	17	AL758979	AL758979 Arabidops
	43	12.6	63.0	67	12	BG892810	BG892810 de56e02.y
	44	12.6	63.0	68	17	AZ660373	AZ660373 1M0538K22
c	45	12.6	63.0	69	9	AA215059	AA215059 mu97e04.r

ALIGNMENTS

RESULT 1
AI198185
LOCUS AI198185 100 bp mRNA linear EST 02-DEC-1998
DEFINITION qi51f03.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:1860029 3', mRNA sequence.
ACCESSION AI198185
VERSION AI198185.1 GI:3750791
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 100)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```
            Insert Length: 1236   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 82.
FEATURES              Location/Qualifiers
     source           1. .100
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1860029"
                      /clone_lib="NCI_CGAP_Brn25"
                      /tissue_type="anaplastic oligodendroglioma"
                      /lab_host="DH10B"
                      /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTT
                      T 3']; double-stranded cDNA was ligated to Eco RI
                      adaptors (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of the modified pT7T3 vector.
                      Library is normalized, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."
BASE COUNT       35 a     13 c     31 g     21 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 9;  Length 100;
  Best Local Similarity   88.9%;  Pred. No. 3.6e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGT 18
           |||||| ||||||||| |
Db       6 TGAGGTCCGGGTGGAACT 23


RESULT 2
BH790276
LOCUS       BH790276                 75 bp    DNA     linear   GSS 02-APR-2002
DEFINITION  SALK_056721.50.40.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_056721.50.40.x, DNA
            sequence.
ACCESSION   BH790276
VERSION     BH790276.1  GI:19883374
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES              Location/Qualifiers
     source           1. .75
                      /organism="Arabidopsis thaliana"
                      /strain="Columbia 0"
                      /db_xref="taxon:3702"
                      /clone="SALK_056721.50.40.x"
                      /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                      /note="PCR was performed on Arabidopsis thaliana lines
                      each of which contains one or more TDNA insertion
                      elements. The resultant fragment for each line was
                      directly sequenced to determine the genomic sequence at
                      the site of insertion. Details of the protocols used can
                      be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       19 a     11 c     25 g     20 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 17;  Length 75;
  Best Local Similarity   93.8%;  Pred. No. 5.2e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GGTGCGGGTGGAAGTG 19
           |||||||||| ||||
Db      45 GGTGCGGGTGGTAGTG 60


RESULT 3
BF507253
LOCUS       BF507253                 61 bp    mRNA    linear   EST 07-DEC-2000
DEFINITION  5923P-29 Pooled green leaf and root tissue Sorghum bicolor cDNA
            clone 5923P-29, mRNA sequence.
ACCESSION   BF507253
VERSION     BF507253.1  GI:11590551
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   Childs,K.L., Klein,R.R., Klein,P.E., Morishige,D.T. and Mullet,J.E.
  TITLE     Mapping Genes on an Integrated Sorghum Genetic and Physical Map
            Using cDNA Selection Technology
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Kevin Childs
            Department of Biochemistry and Biophysics
            Texas A&M University
            College Station, TX 77843, USA
            Tel: 979 845 0832
            Fax: 979 862 4718
            Email: kchilds@unix.tamu.edu.
FEATURES              Location/Qualifiers
     source           1. .61
                      /organism="Sorghum bicolor"
                      /cultivar="BTx623"
                      /db_xref="taxon:4558"
                      /clone="5923P-29"
                      /clone_lib="Pooled green leaf and root tissue"
                      /tissue_type="green leaf and root tissue"
                      /note="Vector: pBluescript II (SK); Site_1: EcoRI; Site_2:
                      EcoRI"
BASE COUNT       12 a      6 c     29 g     14 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 12;  Length 61;
  Best Local Similarity   84.2%;  Pred. No. 6.2e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTG 19
           || |||| | |||||||||
Db      13 TGTGGTGGGTGTGGAAGTG 31


RESULT 4
BE130245
LOCUS       BE130245                 73 bp    mRNA    linear   EST 20-FEB-2001
DEFINITION  L48-385T3 Ice plant Lambda Uni-Zap XR expression library, 48 hours
            NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-385 5',
            mRNA sequence.
ACCESSION   BE130245
VERSION     BE130245.1  GI:8577608
KEYWORDS    EST.
SOURCE      common iceplant.
  ORGANISM  Mesembryanthemum crystallinum
```

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Cushman, J.C.
  TITLE     An expressed sequence tag database for the common ice plant,
            Mesembryanthemum crystallinum
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Cushman JC
            Department of Biochemistry
            University of Nevada
            MS200, Reno, NV 89557-0014, USA
            Tel: 775-784-1918
            Fax: 775-784-1650
            Email: jcushman@unr.edu
            PCR PRimers
            FORWARD: T7
            BACKWARD: T3
            Plate: L48-4  row: H  column: 1
            Seq primer: T3
            High quality sequence stop: 73
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Mesembryanthemum crystallinum"
                     /db_xref="taxon:3544"
                     /clone="L48-385"
                     /clone_lib="Ice plant Lambda Uni-Zap XR expression library
                     , 48 hours NaCl treatment"
                     /tissue_type="Leaf, 48 h 0.4M NaCl"
                     /dev_stage="Six week old"
                     /note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
BASE COUNT       16 a      6 c     30 g     21 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 73;
  Best Local Similarity   84.2%;  Pred. No. 6.2e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTG 19
           ||| ||| |||||||| ||
Db      42 TGAAGTGAGGGTGGAATTG 60


RESULT 5
AA646813/c
LOCUS       AA646813                 82 bp    mRNA    linear   EST 28-OCT-1997
DEFINITION  vn37b09.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:1023353 5' similar to SW:ANXA_RABIT P33477 ANNEXIN XI ;, mRNA
            sequence.
ACCESSION   AA646813
VERSION     AA646813.1  GI:2573242
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:574129
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28ml3 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1023353"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a      25 c     18 g     20 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 82;
  Best Local Similarity   84.2%;  Pred. No. 6.2e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTG 19
           ||||| || | ||||||||
Db      41 TGAGGAGCCGCTGGAAGTG 23


RESULT 6
BG402252
LOCUS       BG402252                 86 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  602465822F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4593941 5',
            mRNA sequence.
ACCESSION   BG402252
VERSION     BG402252.1  GI:13295700
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1334  row: l  column: 06
            High quality sequence stop: 42.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4593941"
                     /clone_lib="NIH_MGC_75"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
```

5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
BASE COUNT 20 a 11 c 23 g 32 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 12; Length 86;
Best Local Similarity 84.2%; Pred. No. 6.2e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGAGGTGCGGGTGGAAGTG 19
||||||| |||||||||
Db 45 TGAGGTGTATGTGGAAGTG 63

RESULT 7
AI559239/c
LOCUS AI559239 99 bp mRNA linear EST 24-MAR-1999
DEFINITION tq32e03.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2210524 3' similar to gb:X77738_rna1 BAND 3 ANION TRANSPORT PROTEIN (HUMAN);, mRNA sequence.
ACCESSION AI559239
VERSION AI559239.1 GI:4509444
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 99)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .99
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2210524"
/clone_lib="NCI_CGAP_Ut1"
/tissue_type="well-differentiated endometrial adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"
BASE COUNT 12 a 25 c 30 g 31 t 1 others
ORIGIN

Query Match 71.0%; Score 14.2; DB 9; Length 99;
Best Local Similarity 84.2%; Pred. No. 6.2e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GAGGTGCGGGTGGAAGTGC 20
|||||| ||| |||||||
Db 88 GAGGTGGAGGTTGAAGTGC 70

RESULT 8
AU102737
LOCUS AU102737 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU102737 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HEP13330, mRNA sequence.
ACCESSION AU102737
VERSION AU102737.1 GI:13552258
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP13330"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and dimethylfumarate treated U937 cells"
BASE COUNT 7 a 14 c 19 g 10 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 50;
Best Local Similarity 88.2%; Pred. No. 8.8e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AGGTGCGGGTGGAAGTG 19
| ||||||| |||||||
Db 1 ACGTGCGGGAGGAAGTG 17

RESULT 9
AZ982649
LOCUS AZ982649 60 bp DNA linear GSS 27-APR-2001
DEFINITION 2M0263H04R Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0263H04 R, DNA sequence.
ACCESSION AZ982649
VERSION AZ982649.1 GI:13853876
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 60)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

```
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0263  row: H  column: 04
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 60.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0263H04"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       15 a      5 c     27 g     13 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 60;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTGC 20
           || |||| ||||||| | ||
Db      40 TGGGGTGAGGGTGGAGGGGC 59


RESULT 10
AA659447
LOCUS       AA659447                 61 bp    mRNA    linear   EST 05-NOV-1997
DEFINITION  nu24b08.s1 NCI_CGAP_Pr2 Homo sapiens cDNA clone IMAGE:1208919
            similar to SW:R13A_HUMAN P40429 60S RIBOSOMAL PROTEIN L13A ;, mRNA
            sequence.
ACCESSION   AA659447
VERSION     AA659447.1  GI:2595601
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1208919"
                     /clone_lib="NCI_CGAP_Pr2"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated,. total cellular RNA obtained from 5,000-10
                     ,000 microdissected preneoplastic cells
                     histologically-determined to be prostatic intraepithelial
                     neoplasia 2 (PIN2) cells.  Double-stranded cDNA was
                     ligated to EcoRI adaptors, 5 cycles of PCR applied to the
                     cDNA with an adaptor-specific primer, and the resulting
                     PCR product subcloned into pAMP10 by the UDG-cloning
                     method (Life Technologies).  Average insert size is 600
                     bp. NOTE: Not directionally cloned.  This library was
                     constructed by  David Krizman."
BASE COUNT        5 a      15 c     25 g     16 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 61;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTGC 20
           |||||||| ||| |  ||||
Db       1 TGAGGTGCAGGTCGTGGTGC 20


RESULT 11
DR17P21S
LOCUS       DR17P21S                 69 bp    DNA     linear   GSS 06-JUN-2002
DEFINITION  Danio rerio genomic clone DKey-17P21, genomic survey sequence.
ACCESSION   AL740955
VERSION     AL740955.1  GI:21343019
KEYWORDS    GSS.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Humphray,S.J., Huckle,E. and Hunt,S.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
            humquery@sanger.ac.uk Unpublished
COMMENT     This sequence was generated from the SP6 end of BAC 17P21. 17P21 is
            part of the Daniokey Pilot BAC Library created by R. Plasterk and
            N.V. Keygene.
            Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="DKey-17P21"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
BASE COUNT       14 a      11 c     22 g     22 t
ORIGIN
```

```
  Query Match             68.0%;  Score 13.6;  DB 17;  Length 69;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGAGGTGCGGGTGGAAGTGC 20
          |||| ||| | |||||| ||
Db     18 TGAGTTGCAGCTGGAAGCGC 37


RESULT 12
AA500773
LOCUS       AA500773                 88 bp    mRNA    linear   EST 01-JUL-1997
DEFINITION  vg01b07.r1 Soares mouse NbMH Mus musculus cDNA clone IMAGE:860053
            5' similar to WP:F42H10.6 CE00168 ;, mRNA sequence.
ACCESSION   AA500773
VERSION     AA500773.1  GI:2235740
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:504141
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .88
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:860053"
                     /clone_lib="Soares mouse NbMH"
                     /sex="male"
                     /tissue_type="heart"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAAAGTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Minoru Ko, Wayne State Univ. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
BASE COUNT       26 a     13 c     27 g     22 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 88;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGAGGTGCGGGTGGAAGTGC 20
          |||| ||  ||||||||| ||
Db     58 TGAGATGAAGGTGGAAGAGC 77


RESULT 13
H55398/c
LOCUS       H55398                   92 bp    mRNA    linear   EST 07-DEC-1995
DEFINITION  CHR220337 Chromosome 22 exon Homo sapiens cDNA clone C22_433 5',
            mRNA sequence.
ACCESSION   H55398
VERSION     H55398.1  GI:1108264
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Trofatter,J.A., Long,K.R., Murrell,J.R., Stotler,C.J., Gusella,J.F.
            and Buckler,A.J.
  TITLE     An expression-independent catalog of genes from human chromosome 22
  JOURNAL   Genome Res. 5 (3), 214-224 (1995)
  MEDLINE   96159527
COMMENT     Contact: Buckler AJ
            Molecular Neurogenetics Unit
            Massachusetts General Hospital
            Building 149, 13th St., Charlestown MA 02129
            Tel: 6177249616
            Fax: 6177265736
            Email: buckler@helix.mgh.harvard.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="C22_433"
                     /clone_lib="Chromosome 22 exon"
                     /lab_host="E. coli DH5a"
                     /note="Vector: pBluescriptIIKS+; Site_1: Sal I; Site_2:
                     Bam HI (destroyed); Exons were isolated from human
                     chromosome 22 specific cosmids using a modification of
                     the method of exon amplification (Proc. Natl. Acad. Sci.
                     USA 88:4005-4009, 1991). Amplified exons were digested
                     with Sal I and Bgl II and subsequently cloned into
                     pBluescriptIIKS+ at the Sal I and Bam HI sites."
BASE COUNT       18 a     28 c     32 g     14 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 92;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGAGGTGCGGGTGGAAGTGC 20
          |||  |||  ||||||||||
Db     55 TGACCTGCCCGTGGAAGTGC 36


RESULT 14
Z20250/c
LOCUS       Z20250                   96 bp    mRNA    linear   EST 07-FEB-1995
DEFINITION  HSAAABRWI P, Human foetal Brain Whole tissue Homo sapiens cDNA,
            mRNA sequence.
ACCESSION   Z20250
VERSION     Z20250.1  GI:26995
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   MRC Human Genome Mapping Project Resource Centre.
  TITLE     The UK-HGMP cDNA program
  JOURNAL   Unpublished (1993)
COMMENT     Contact: MRC Human Genome Mapping Project Resource Centre
```

```
            Clinical Research Centre
            Watford Road,Harrow,Middlesex HA1 3UJ,U.K.
            Email: biohelp@hgmp.mrc.ac.uk
            single read.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="P, Human foetal Brain Whole tissue"
                     /note="Vector: Bluescript; clone_library=P, Human foetal
                     Brain Whole tissue; cloning vector is Bluescript. "
BASE COUNT       26 a     29 c     22 g     17 t      2 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 96;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGAGGTGCGGGTGGAAGTGC 20
           || |||| | ||||||| ||
Db      40 TGGGGTGGGCGTGGAAGGGC 21


RESULT 15
AQ248436/c
LOCUS       AQ248436                 72 bp    DNA     linear   GSS 06-OCT-1998
DEFINITION  T7M22-Sp6 TAMU Arabidopsis thaliana genomic clone T7M22, DNA
            sequence.
ACCESSION   AQ248436
VERSION     AQ248436.1  GI:3698519
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
            Ecker,J.
  TITLE     BAC End Sequences at ATGC
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Ecker J.
            Arabidopsis Thaliana Genome Center
            University of Pennsylvania
            Dept. of Biology, University of Pennsylvania, Philadelphia, PA
            19104
            Tel: 215-898-9384
            Fax: 215-898-8780
            Email: jecker@atgenome.bio.upenn.edu
            Seq primer: Sp6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="T7M22"
                     /clone_lib="TAMU"
                     /sex="hermaphrodite"
                     /note="Vector: BeloBACII; Site_1: HindIII; Site_2: HindIII
                     ; Produced by Rod Wing"
BASE COUNT        8 a     39 c      4 g     21 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 72;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTG 19
           ||||| ||||||| || |
Db      61 GAGGTACGGGTGGTAGGG 44
```

Search completed: November 26, 2002, 04:07:48
Job time : 766.8 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-10
Perfect score: 20
Sequence: 1 gtgccgacgtgggacccaga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	17.6	88.0	21	6	AX097265	AX097265 Sequence
2	13.8	69.0	27	6	AX188785	AX188785 Sequence
3	13.6	68.0	27	6	A98650	A98650 Sequence 6
c 4	13.6	68.0	51	6	AX162818	AX162818 Sequence
c 5	13.2	66.0	27	6	AX002713	AX002713 Sequence
6	13.2	66.0	33	6	AR127237	AR127237 Sequence
c 7	13.2	66.0	33	6	AR127238	AR127238 Sequence
c 8	13.2	66.0	35	6	AX358674	AX358674 Sequence
9	13.2	66.0	35	6	AX358675	AX358675 Sequence
c 10	13.2	66.0	44	6	AR143578	AR143578 Sequence
c 11	13.2	66.0	44	6	AR168947	AR168947 Sequence
12	13.2	66.0	90	9	HSEXTR2	X99574 H.sapiens e
13	12.8	64.0	22	6	E11187	E11187 Probe. 9/19
c 14	12.8	64.0	24	6	AX019309	AX019309 Sequence
15	12.8	64.0	25	6	AR065111	AR065111 Sequence
16	12.8	64.0	25	6	E09880	E09880 Primer. 9/1
c 17	12.8	64.0	30	9	HUMPLTP12	U37819 Human phosp
c 18	12.8	64.0	48	6	A97298	A97298 Sequence 15
19	12.8	64.0	53	6	E11189	E11189 Probe. 9/19
c 20	12.8	64.0	62	6	AR127979	AR127979 Sequence
c 21	12.8	64.0	70	6	E03982	E03982 Probe for d
22	12.8	64.0	100	10	RATMAL01	M26581 Rattus norv
c 23	12.6	63.0	20	6	AR121581	AR121581 Sequence
c 24	12.6	63.0	30	6	AX173016	AX173016 Sequence
c 25	12.6	63.0	30	6	AX173171	AX173171 Sequence
c 26	12.6	63.0	30	6	AX174698	AX174698 Sequence
c 27	12.6	63.0	45	6	AR168071	AR168071 Sequence
c 28	12.6	63.0	45	6	AR204842	AR204842 Sequence
c 29	12.6	63.0	51	6	AX162807	AX162807 Sequence
c 30	12.6	63.0	51	6	AX162817	AX162817 Sequence
c 31	12.6	63.0	51	6	AX162819	AX162819 Sequence
c 32	12.6	63.0	51	6	AX162820	AX162820 Sequence
33	12.6	63.0	83	9	HSRNPT2S15	U56689 Human renal
c 34	12.6	63.0	84	10	S56855	S56855 GH-releasin
c 35	12.6	63.0	92	9	HUMPK30	D28371 Human mRNA
36	12.4	62.0	50	6	AX159092	AX159092 Sequence
37	12.4	62.0	50	6	AX427051	AX427051 Sequence
38	12.4	62.0	51	6	AX427050	AX427050 Sequence
39	12.4	62.0	80	10	MMDNDS9	X55227 M.musculus
c 40	12.2	61.0	19	6	AX132262	AX132262 Sequence
c 41	12.2	61.0	20	6	AR206645	AR206645 Sequence
c 42	12.2	61.0	20	6	AR206646	AR206646 Sequence
43	12.2	61.0	23	6	AR137741	AR137741 Sequence
44	12.2	61.0	30	6	A95831	A95831 Sequence 4
45	12.2	61.0	34	6	AR108658	AR108658 Sequence

ALIGNMENTS

RESULT 1
AX097265/c
LOCUS AX097265 21 bp DNA linear PAT 30-MAR-2001
DEFINITION Sequence 2443 from Patent WO0118250.
ACCESSION AX097265
VERSION AX097265.1 GI:13513682
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 21)
AUTHORS Lander,E.S., Gargill,M., Ireland,J.S., Bolk,S., Daley,G.Q. and
Mccarthy,J.J.
TITLE Single nucleotide polymorphisms in genes

```
  JOURNAL   Patent: WO 0118250-A 2443 15-MAR-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium
            Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        3 a      8 c      5 g      4 t      1 others
ORIGIN

  Query Match             88.0%;  Score 17.6;  DB 6;  Length 21;
  Best Local Similarity   94.4%;  Pred. No. 1.8e+02;
  Matches   17;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCA 18
           |||||||:||||||||||
Db      18 GTGCCGAYGTGGGACCCA 1


RESULT 2
AX188785
LOCUS       AX188785                  27 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 32 from Patent WO0148228.
ACCESSION   AX188785
VERSION     AX188785.1  GI:15142326
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Chater,K.F., Bruton,C.J., O'Rouke,S.J. and Wietzorrek,A.W.
  TITLE     Methods and materials relating to gene expression
  JOURNAL   Patent: WO 0148228-A 32 05-JUL-2001;
            Plant Bioscience Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Oligonucleotide primer"
BASE COUNT        3 a      12 c      9 g      3 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 27;
  Best Local Similarity   88.2%;  Pred. No. 1.7e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GCCGACGTGGGACCCAG 19
           ||||||| ||||||| |
Db       3 GCCGACGCGGGACCCCG 19


RESULT 3
A98650
LOCUS       A98650                    27 bp    DNA     linear   PAT 26-JAN-2000
DEFINITION  Sequence 6 from Patent WO9910500.
ACCESSION   A98650
VERSION     A98650.1  GI:6781694
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Schlichter,U. and Steinbiss,H.
  TITLE     NUCLEIC ACID MOLECULES CODING FOR A CYSTEINE PROTEINASE OF PLANT
            ORIGIN AND THEIR REGULATING REGIONS
  JOURNAL   Patent: WO 9910500-A 6 04-MAR-1999;
            MAX PLANCK GESELLSCHAFT (DE); SCHLICHTER URSULA (DE)
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        6 a      6 c     11 g      4 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 27;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCAGA 20
           |||  ||| || ||||||||
Db       6 GTGTGGACCTGCGACCCAGA 25


RESULT 4
AX162818/c
LOCUS       AX162818                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 6146 from Patent WO0140521.
ACCESSION   AX162818
VERSION     AX162818.1  GI:14544149
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 6146 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (6145 is other entry)
                     Accession number cg44914955"
BASE COUNT       14 a      18 c      12 g      7 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCCGACGTGGGACCCAGA 20
           || | |||||||| | ||||
Db      45 GTTCAGACGTGGGTCGCAGA 26


RESULT 5
AX002713/c
LOCUS       AX002713                  27 bp    DNA     linear   PAT 21-AUG-2000
DEFINITION  Sequence 5 from Patent WO9900490.
ACCESSION   AX002713
VERSION     AX002713.1  GI:9885040
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Wainberg,M.A.
  TITLE     Attenuated human immunodeficiency virus vaccine
  JOURNAL   Patent: WO 9900490-A 5 07-JAN-1999;
            WAINBERG MARK A (CA); SIR MORTIMER B DAVIS JEWISH GE (CA)
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        3 a      7 c     11 g      6 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 27;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCCGACGTGGGACCCAGA 20
```

```
RESULT 6
AF090124
LOCUS       AF090124                 78 bp    mRNA    linear   INV 07-OCT-1998
DEFINITION  Helix aspersa shaw-like potassium channel Kv3.1 (Kv3.1) mRNA,
            partial cds.
ACCESSION   AF090124
VERSION     AF090124.1  GI:3703091
KEYWORDS    .
SOURCE      Helix aspersa.
  ORGANISM  Helix aspersa
            Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata;
            Stylommatophora; Helicacea; Helicidae; Helix.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Witchel,H.J., Davies,D. and Meech,R.W.
  TITLE     Immunological identification of a putative potassium channel
            subunit from a short (9) amino acid sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 78)
  AUTHORS   Witchel,H.J., Davies,D. and Meech,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-SEP-1998) Physiology, Medical School, University of
            Bristol, Bristol BS8 1TD, England
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Helix aspersa"
                     /db_xref="taxon:6535"
                     /tissue_type="brain"
     gene            <1. .>78
                     /gene="Kv3.1"
     CDS             <1. .>78
                     /gene="Kv3.1"
                     /codon_start=1
                     /product="shaw-like potassium channel Kv3.1"
                     /protein_id="AAC63041.1"
                     /db_xref="GI:3703092"
                     /translation="DMVPKTYVGMFVGALCALAGVLCIAL"
BASE COUNT       10 a     17 c     28 g     23 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 3;  Length 78;
  Best Local Similarity   93.3%;  Pred. No. 5.1e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 TGGCGCCCTGTGTCC 17
           ||||||||||||| |
Db      36 TGGCGCCCTGTGTGC 50


RESULT 7
AR163949/c
LOCUS       AR163949                 20 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 147 from patent US 6271030.
ACCESSION   AR163949
VERSION     AR163949.1  GI:16234807
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Monia,B.P., Butler,M.M. and Wyatt,J.
  TITLE     Antisense inhibition of C/EBP beta expression
  JOURNAL   Patent: US 6271030-A 147 07-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        6 a      7 c      6 g      1 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 7.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATGGCGCCCTGTGTCCCG 19
           |||||| || || |||||
Db      20 ATGGCGGCCGGTTTCCCG 3


RESULT 8
AX326934/c
LOCUS       AX326934                 20 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 130 from Patent WO0178894.
ACCESSION   AX326934
VERSION     AX326934.1  GI:18097645
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Keith,T.
  TITLE     Novel human gene relating to respiratory diseases, obesity, and
            inflammatory bowel disease
  JOURNAL   Patent: WO 0178894-A 130 25-OCT-2001;
            Genome Therapeutics Corp. (US)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer"
BASE COUNT        7 a      3 c      7 g      3 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 7.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 AATGGCGCCCTGTGTCCC 18
           |||||  |||| ||||||
Db      18 AATGGTTCCCTCTGTCCC 1


RESULT 9
A60217
LOCUS       A60217                   34 bp    DNA     linear   PAT 06-MAR-1998
DEFINITION  Sequence 13 from Patent WO9708330.
ACCESSION   A60217
VERSION     A60217.1  GI:3715225
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.
  TITLE     EXPRESSION SYSTEMS
  JOURNAL   Patent: WO 9708330-A 13 06-MAR-1997;
            CANCER RES CAMPAIGN TECH (GB)
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        6 a     10 c      9 g      9 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 34;
  Best Local Similarity   83.3%;  Pred. No. 6.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATGGCGCCCTGTGTCCCG 19
           |||||||||| || | ||
Db       5 ATGGCGCCCTATGGCTCG 22


RESULT 10
AR122293
LOCUS       AR122293                 34 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 13 from patent US 6165715.
ACCESSION   AR122293
```

```
VERSION     AR122293.1  GI:14106610
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Collins,M.Katherine.Levinge., Weiss,R.Anthony., Takeuchi,Y. and
            Cosset,F.-L.
  TITLE     Expression systems
  JOURNAL   Patent: US 6165715-A 13 26-DEC-2000;
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="unknown"
BASE COUNT        6 a      10 c       9 g       9 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 34;
  Best Local Similarity   83.3%;  Pred. No. 6.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATGGCGCCCTGTGTCCCG 19
           |||||||||| || | ||
Db       5 ATGGCGCCCTATGGCTCG 22


RESULT 11
AR135209/c
LOCUS       AR135209                 40 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 35 from patent US 6194559.
ACCESSION   AR135209
VERSION     AR135209.1  GI:14124114
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Kim,S.Young.
  TITLE     Abscisic acid responsive element-binding transcription factors
  JOURNAL   Patent: US 6194559-A 35 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="unknown"
BASE COUNT        9 a      11 c      14 g       6 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 6.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGGCGCCCTGTGTCCCGA 20
           | |||||  |||||||||
Db      26 TTGCGCCACGTGTCCCGA 9


RESULT 12
AR146705/c
LOCUS       AR146705                 40 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 35 from patent US 6218527.
ACCESSION   AR146705
VERSION     AR146705.1  GI:15109894
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Kim,S.Young.
  TITLE     Nucleic acid molecule encoding abscisic acid responsive
            element-binding factor 3
  JOURNAL   Patent: US 6218527-A 35 17-APR-2001;
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="unknown"
BASE COUNT        9 a      11 c      14 g       6 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 6.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGGCGCCCTGTGTCCCGA 20
           | |||||  |||||||||
Db      26 TTGCGCCACGTGTCCCGA 9


RESULT 13
AR152276/c
LOCUS       AR152276                 40 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 35 from patent US 6232461.
ACCESSION   AR152276
VERSION     AR152276.1  GI:15118326
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Kim,S.Young.
  TITLE     Nucleic acid molecule encoding abscisic acid responsive
            element-binding ractor 4
  JOURNAL   Patent: US 6232461-A 35 15-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="unknown"
BASE COUNT        9 a      11 c      14 g       6 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 6.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGGCGCCCTGTGTCCCGA 20
           | |||||  |||||||||
Db      26 TTGCGCCACGTGTCCCGA 9


RESULT 14
AR157814/c
LOCUS       AR157814                 40 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 35 from patent US 6245905.
ACCESSION   AR157814
VERSION     AR157814.1  GI:16218827
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Kim,S.Young.
  TITLE     Nucleic acid molecule encoding abscisic acid responsive
            element-binding factor 2
  JOURNAL   Patent: US 6245905-A 35 12-JUN-2001;
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="unknown"
BASE COUNT        9 a      11 c      14 g       6 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 6.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGGCGCCCTGTGTCCCGA 20
           | |||||  |||||||||
Db      26 TTGCGCCACGTGTCCCGA 9


RESULT 15
HSA403877
```

```
LOCUS       HSA403877                45 bp    mRNA    linear   PRI 24-MAY-2000
DEFINITION  Homo sapiens partial mRNA for histone specific T-cell receptor beta
            chain, V-D-J region (TCRBV12.1-J1.2 gene).
ACCESSION   AJ403877
VERSION     AJ403877.1  GI:8217172
KEYWORDS    diversity region; joining segment; T-cell receptor beta; TCR beta
            chain; TCRBV12.1-J1.2 gene; variable region.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Andreassen,K., Bendiksen,S., Moens,U., Van Ghelue,M. and Rekvig,O.
  TITLE     Comparison of precursor frequencies and receptor structures of
            histone specific T cells from SLE patients and from normal
            individuals reveals no major differences
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 45)
  AUTHORS   Andreassen,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-2000) Andreassen K., Molecular Genetics, Medical
            Biology, University of Tromsoe, Mh, 9037 Tromsoe, NORWAY
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Homo sapiens"
                     /isolate="Normal individual N3"
                     /db_xref="taxon:9606"
                     /clone="N3H12"
                     /cell_type="T lymphocyte"
                     /rearranged
     gene            complement(1. .45)
                     /gene="TCRBV12.1-J1.2"
     CDS             complement(<1. .>45)
                     /gene="TCRBV12.1-J1.2"
                     /function="histone specific"
                     /note="V-D-J region"
                     /codon_start=1
                     /product="T-cell receptor beta chain"
                     /protein_id="CAB92457.1"
                     /db_xref="GI:8217173"
                     /translation="CAIGSQGARGYTFGS"
BASE COUNT       10 a     16 c     12 g      7 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 45;
  Best Local Similarity   83.3%;  Pred. No. 6.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGGCGCCCTGTGTCCCGA 20
           | || ||||||| |||||
Db      21 TAGCCCCCTGTGACCCGA 38


Search completed: December  3, 2002, 18:12:42
Job time : 310.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-7
Perfect score: 20
Sequence: 1 aatggcgccctgtgtcccga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31437	Human neuropilin m
	2	14.2	71.0	25	16	AAQ92990	Pre-invasive human
	3	13.8	69.0	41	22	AAH44307	Human fibrinogen 9
	4	13.8	69.0	44	21	AAA47032	PCR primer used to
c	5	13.8	69.0	45	21	AAA47033	PCR primer used to
c	6	13.6	68.0	30	22	AAD21726	Alternative versio
	7	13.6	68.0	47	21	AAZ69022	Human map-related
c	8	13.6	68.0	65	24	ABN58471	Mouse spliced tran
c	9	13.4	67.0	31	22	AAI30331	Human single nucle
	10	13.4	67.0	33	22	AAH44305	Human fibrinogen 9
	11	13.4	67.0	90	24	ABK36795	Human DNA encoding
c	12	13.2	66.0	20	24	ABL94381	Mouse C/EBP beta p
	13	13.2	66.0	34	18	AAT90699	Primer #4 for env
c	14	13.2	66.0	40	22	AAD14556	Arabidopsis thalia
c	15	13.2	66.0	40	22	AAS07861	Binding site selec
c	16	13.2	66.0	40	22	AAD08929	Arabidopsis thalia
c	17	13.2	66.0	40	22	AAS03341	A group IA abscisi
c	18	13.2	66.0	40	22	AAS00394	ABRE binding seque
c	19	13.2	66.0	48	15	AAQ58903	Human beta-glucuro
	20	12.8	64.0	21	13	AAQ22357	DNA for modulating
c	21	12.8	64.0	21	13	AAQ22382	DNA for modulating
c	22	12.8	64.0	21	16	AAT11952	Antisense oligonuc
	23	12.8	64.0	21	16	AAT12015	CMV DNA polymerase
c	24	12.8	64.0	21	16	AAT01631	Peptide nucleic ac
c	25	12.8	64.0	21	19	AAX17869	Anti-CMV oligonucl
	26	12.8	64.0	21	19	AAX17932	CMV target sequenc
	27	12.8	64.0	31	16	AAT01688	Peptide nucleic ac
	28	12.8	64.0	31	19	AAV67700	Nucleotide fragmen
	29	12.8	64.0	31	20	AAX39267	Human genomic DNA
	30	12.8	64.0	31	20	AAX39114	Human genomic DNA
c	31	12.8	64.0	53	19	AAV70985	Primer 5'GFPc used
c	32	12.8	64.0	53	21	AAD00045	5' GFP PCR primer
c	33	12.8	64.0	57	18	AAT70206	RNA polymerase II
c	34	12.8	64.0	60	24	ABN43652	Human spliced tran
c	35	12.6	63.0	27	21	AAA71586	Human MPROT13 prob
c	36	12.6	63.0	60	24	ABN37294	Human spliced tran
	37	12.6	63.0	60	24	ABN40783	Human spliced tran
c	38	12.6	63.0	60	24	ABN44841	Human spliced tran
c	39	12.6	63.0	65	24	ABN52637	Mouse spliced tran
	40	12.6	63.0	72	18	AAT68554	M6P/IGF-II recepto
c	41	12.6	63.0	90	24	ABK36689	Human DNA encoding
	42	12.6	63.0	93	13	AAQ27938	93 mer of human be
c	43	12.4	62.0	41	22	AAH73403	Human G protein co
	44	12.4	62.0	51	22	AAL27440	Human SNP oligonuc
	45	12.4	62.0	51	22	AAL28285	Human SNP oligonuc

ALIGNMENTS

RESULT 1
AAZ31437
ID AAZ31437 standard; DNA; 20 BP.
XX
AC AAZ31437;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3607.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 3 A; 7 C; 6 G; 4 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AATGGCGCCCTGTGTCCCGA 20
           ||||||||||||||||||||
Db       1 AATGGCGCCCTGTGTCCCGA 20


RESULT 2
AAQ92990
ID   AAQ92990 standard; DNA; 25 BP.
XX
AC   AAQ92990;
XX
DT   02-APR-1996  (first entry)
XX
DE   Pre-invasive human breast cancer marker gene PCR primer.
XX
KW   BRCA1; breast cancer; diagnosis; prognosis; gene therapy;
KW   non-comedo DCIS; ductal carcinoma in situ; intraductal carcinoma;
KW   pre-invasive human breast tissue; marker gene; RT-PCR; primer;
KW   randomly selected; ss.
XX
OS   Synthetic.
XX
PN   WO9519369-A1.
XX
PD   20-JUL-1995.
XX
PF   17-JAN-1995;   95WO-US00608.
XX
PR   17-JAN-1995;   95US-0373799.
PR   14-JAN-1994;   94US-0182961.
XX
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Holt JT, Jensen RA, Obermiller PS, Page DL, Robinson-Benion CL;
PI   Thompson ME;
XX
DR   WPI; 1995-269208/35.
XX
PT   Detection, diagnosis and treatment of pre-invasive breast cancer -
PT   by identifying differentially expressed marker genes, also use of
PT   BRCA1 gene in therapy of breast cancer.
XX
PS   Claim 13; Page 106; 149pp; English.
XX
CC   In a novel method, differentially expressed cDNA clones are
CC   identified by comparing cDNA obtained from abnormal breast tissue
CC   (e.g. ductal carcinoma in situ (DCIS)) samples with those obtained
CC   from normal breast epithelial cells. Such clones are useful as
CC   marker genes for pre-invasive human breast tissue. In a prefd.
CC   version of the method, differential expression of the marker gene
CC   is confirmed by using PCR amplification. The present sequence is
CC   that of a randomly selected PCR primer for use in the amplification.
XX
SQ   Sequence 25 BP; 4 A; 8 C; 8 G; 5 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 16;  Length 25;
  Best Local Similarity   84.2%;  Pred. No. 1.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATGGCGCCCTGTGTCCCGA 20
           |||||| |||||  |||||
Db       3 ATGGCGTCCTGTAACCCGA 21


RESULT 3
AAH44307
ID   AAH44307 standard; DNA; 41 BP.
XX
AC   AAH44307;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human fibrinogen 9 probe SEQ ID NO:8.
XX
KW   Human; fibrinogen 9; cytostatic; haemostatic; antiinflammatory;
KW   immunomodulatory; virucide; malignant tumour; haemopathy;
KW   HIV infection; immunological disease; inflammation; probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200148196-A1.
XX
PD   05-JUL-2001.
XX
PF   18-DEC-2000; 2000WO-CN00611.
XX
PR   24-DEC-1999;   99CN-0125762.
XX
PA   (BIOW-) BIOWINDOW GENE DEV LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2001-418278/44.
XX
PT   Fibrinogen 9 and encoded polynucleotide, applicable in diagnosis and
PT   treatment of malignant tumor, hemopathy, HIV infection, immunological
PT   diseases and various inflammations -
XX
PS   Example 7; Page 14; 27pp; Chinese.
XX
CC   The present sequence represents a probe for human fibrinogen 9.
CC   Human fibrinogen 9 has cytostatic, haemostatic, antiinflammatory,
CC   immunomodulatory and virucide activities. The human fibrinogen 9 protein
CC   and polynucleotide sequences are applicable in the diagnosis and
CC   treatment of malignant tumour, haemopathy, HIV infection, immunological
CC   diseases and various inflammations.
XX
SQ   Sequence 41 BP; 11 A; 9 C; 11 G; 10 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 41;
  Best Local Similarity   88.2%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GGCGCCCTGTGTCCCGA 20
           || |||||||||||| |
Db       8 GGAGCCCTGTGTCCCCA 24


RESULT 4
AAA47032
ID   AAA47032 standard; DNA; 44 BP.
XX
AC   AAA47032;
```

XX
DT 03-OCT-2000 (first entry)
XX
DE PCR primer used to introduce mutations S19A/S23A in TRIP9.
XX
KW Human; IKK3 kinase; IL-8 gene regulator; anti-inflammatory; immunogen;
KW PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200039308-A1.
XX
PD 06-JUL-2000.
XX
PF 24-DEC-1999; 99WO-JP07286.
XX
PR 24-DEC-1998; 98GB-0028704.
XX
PA (GLAX) GLAXO WELLCOME KK.
XX
PI Takemoto Y, Sakai Y, Hashimoto Y;
XX
DR WPI; 2000-475700/41.
XX
PT New IKK3 kinase protein and nucleotides encoding it, useful for
PT screening for IKK3 protein modulators for treating inflammation, e.g.
PT arthritis, atopic dermatitis or systemic lupus erythematosus -
XX
PS Example 1; Page 17; 102pp; English.
XX
CC PCR primers AAA47032-33 were used to modify DNA encoding a human
CC TRIP9. The specification describes an IKK3 protein. The protein
CC is an interleukin-8 (IL-8) gene regulator. The IKK3 kinase protein
CC is useful for screening for agents with anti-inflammatory activity.
CC Compounds which exhibits IKK3 kinase modulating activity may be used
CC in formulating a treatment or prophylaxis of a disorder responsive
CC to the modulation of IKK3 kinase activity in a mammal. The proteins,
CC their variants or fragments, derivatives, analogues or cells
CC expressing them can also be used as immunogens to produce antibodies
CC against IKK3 protein, which may further be used to locate the protein
CC in tissues expressing that protein.
XX
SQ Sequence 44 BP; 6 A; 14 C; 18 G; 6 T; 0 other;

Query Match 69.0%; Score 13.8; DB 21; Length 44;
Best Local Similarity 88.2%; Pred. No. 2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGCGCCCTGTGTCCCGA 20
||||||||| |||| ||
Db 24 GGCGCCCTGGGTCCGGA 40

RESULT 5
AAA47033/c
ID AAA47033 standard; DNA; 45 BP.
XX
AC AAA47033;
XX
DT 03-OCT-2000 (first entry)
XX
DE PCR primer used to introduce mutations S19A/S23A in TRIP9.
XX
KW Human; IKK3 kinase; IL-8 gene regulator; anti-inflammatory; immunogen;
KW PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200039308-A1.
XX
PD 06-JUL-2000.
XX
PF 24-DEC-1999; 99WO-JP07286.
XX
PR 24-DEC-1998; 98GB-0028704.
XX
PA (GLAX) GLAXO WELLCOME KK.
XX
PI Takemoto Y, Sakai Y, Hashimoto Y;
XX
DR WPI; 2000-475700/41.
XX
PT New IKK3 kinase protein and nucleotides encoding it, useful for
PT screening for IKK3 protein modulators for treating inflammation, e.g.
PT arthritis, atopic dermatitis or systemic lupus erythematosus -
XX
PS Example 1; Page 17; 102pp; English.
XX
CC PCR primers AAA47032-33 were used to modify DNA encoding a human
CC TRIP9. The specification describes an IKK3 protein. The protein
CC is an interleukin-8 (IL-8) gene regulator. The IKK3 kinase protein
CC is useful for screening for agents with anti-inflammatory activity.
CC Compounds which exhibits IKK3 kinase modulating activity may be used
CC in formulating a treatment or prophylaxis of a disorder responsive
CC to the modulation of IKK3 kinase activity in a mammal. The proteins,
CC their variants or fragments, derivatives, analogues or cells
CC expressing them can also be used as immunogens to produce antibodies
CC against IKK3 protein, which may further be used to locate the protein
CC in tissues expressing that protein.
XX
SQ Sequence 45 BP; 6 A; 19 C; 14 G; 6 T; 0 other;

Query Match 69.0%; Score 13.8; DB 21; Length 45;
Best Local Similarity 88.2%; Pred. No. 2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGCGCCCTGTGTCCCGA 20
||||||||| |||| ||
Db 21 GGCGCCCTGGGTCCGGA 5

RESULT 6
AAD21726/c
ID AAD21726 standard; DNA; 30 BP.
XX
AC AAD21726;
XX
DT 28-JAN-2002 (first entry)
XX
DE Alternative version of mouse SCR-1 cDNA amplifying PCR primer.
XX
KW Mouse; stem cell growth factor-like protein; antiinflammatory; nootropic;
KW neuroprotective; vulnerary; cytostatic; anticonvulsant; immunostimulant;
KW vasotropic; virucide; dermatological; tranquilliser; cerebroprotective;
KW osteopathic; immunodeficiency syndrome; chronic granulomatous disease;
KW duplicated immunodeficiency syndrome; Wiskott-Aldrich syndrome; AIDS;
KW acquired immune deficiency syndrome; agammaglobulinaemia; thalassaemia;
KW Gaucher's disease; lysosomal storage disease; mucopolysaccharidosis;
KW adrenal white matter degeneration; anaemia; neurodegenerative disease;
KW Parkinson's disease; Alzheimer's disease; thrombocytopaenia; SCID;
KW severe combined immunodeficiency; immune disorder; autoimmune disorder;
KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
KW autoimmune pulmonary inflammation; cancer; tumour; gene therapy; SCR-1;
KW supporting factor for the proliferation of stem cell; PCR primer; ss.
XX
OS Mus musculus.
XX
PN WO200177169-A2.
XX
PD 18-OCT-2001.
XX
PF 05-APR-2001; 2001WO-US11208.
XX
PR 05-APR-2000; 2000US-0543774.
PR 28-JUN-2000; 2000US-215733P.
PR 09-JAN-2001; 2001US-0757562.

```
PR   05-FEB-2001; 2001US-266614P.
XX
PA   (HYSE-) HYSEQ INC.
PA   (KIRI ) KIRIN BEER KK.
XX
PI   Tang TY, Labat I, Tillinghast JS, Sinku A, Liu C, Drmanac RT;
PI   Stache-Crain B, Dickson M, Mize NK, Nishikawa M;
XX
DR   WPI; 2001-657166/75.
XX
PT   Novel stem cell growth factor like polypeptides and polynucleotides for
PT   identifying modulators useful for treating diseases such as Alzheimer's
PT   disease, cancer, rheumatoid arthritis, osteoporosis -
XX
PS   Example 12; Page 221; 232pp; English.
XX
CC   The patent discloses novel stem cell growth factor-like proteins and
CC   polynucleotides encoding them. Proteins of the invention are also known
CC   as supporting factor for the proliferation of stem cells (SCR-1). Stem
CC   cell growth factor-like proteins are useful for supporting proliferation
CC   or survival of a stem cell or germ cell which is preferably primordial
CC   germ cell, germ line stem cell, embryonic stem cell, haematopoietic stem
CC   cell, haematopoietic progenitor cell, pluripotent cell or totipotent
CC   cell. The haematopoietic progenitor cell cultured using stem cell
CC   growth factor-like proteins can replace as a graft for the bone marrow
CC   transplantation or cord blood transplantation for treating a variety
CC   of diseases such as immunodeficiency syndrome, chronic granulomatous
CC   disease, duplicated immunodeficiency syndrome, agammaglobulinaemia,
CC   Wiskott-Aldrich syndrome, acquired immune deficiency syndrome (AIDS),
CC   thalassaemia, haemolytic anaemia due to enzyme defect, congenital
CC   anaemia such as sickle cell anaemia, Gaucher's disease, lysosomal
CC   storage diseases such as mucopolysaccharidosis, adrenal white matter
CC   degeneration, a variety of cancer and tumours. Proteins of the
CC   invention are useful for treating diseases such as Parkinson's
CC   disease, Alzheimer's disease and other neurodegenerative diseases,
CC   thrombocytopaenia, immune deficiencies and disorders such as severe
CC   combined immunodeficiency (SCID)) and autoimmune disorders such as
CC   multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
CC   and autoimmune pulmonary inflammation. Sequences of the invention are
CC   also useful in gene therapy. The present DNA sequence is the alternative
CC   version of mouse SCR-1 cDNA amplifying PCR primer.
CC   Note: This sequence is stated to be the same as that shown as SEQ ID
CC   NO: 30 in page 179 of the specification. However these sequences differ
CC   at many locations.
XX
SQ   Sequence 30 BP; 5 A; 6 C; 11 G; 8 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 22;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 2.5e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AATGGCGCCCTGTGTCCCGA 20
           |||||||| ||   ||||||
Db      21 AATGGCGCGCTTCTTCCCGA 2


RESULT 7
AAZ69022
ID   AAZ69022 standard; DNA; 47 BP.
XX
AC   AAZ69022;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:3378.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation;
KW   diagnosis; single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,T)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB00822.
XX
PR   21-APR-1998;   98US-0082614.
PR   23-NOV-1998;   98US-0109732.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D, Blumenfeld M, Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome -
XX
PS   Claim 3; Page 951; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the
CC   invention have a variety of uses: they can be used for high density
CC   mapping of the human genome, and in complex association studies and
CC   haplotyping studies which are useful in determining the genetic basis
CC   for disease states. Compositions and methods of the invention can also
CC   be useful for the identification of the targets for the development of
CC   pharmaceutical agents and diagnostic methods, as well as the
CC   characterisation of the differential efficacious responses to and side
CC   effects from pharmaceutical agents acting on a disease as well as other
CC   treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC   and 3367, are not actually given a sequence in the Sequence Listing
CC   from the present invention.
XX
SQ   Sequence 47 BP; 6 A; 14 C; 18 G; 9 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 21;  Length 47;
  Best Local Similarity   80.0%;  Pred. No. 2.5e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AATGGCGCCCTGTGTCCCGA 20
           || | || |||||| |||||
Db       7 AAGGCCGTCCTGTGCCCCGA 26


RESULT 8
ABN58471/c
ID   ABN58471 standard; DNA; 65 BP.
XX
AC   ABN58471;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse spliced transcript detection oligonucleotide SEQ ID NO:31219.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Mus musculus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
```

```
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 31219; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 15 A; 21 C; 13 G; 16 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 65;
  Best Local Similarity   80.0%;  Pred. No. 2.6e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AATGGCGCCCTGTGTCCCGA 20
           ||||| ||| | ||||| ||
Db      60 AATGGTGCCTTCTGTCCTGA 41


RESULT 9
AAI30331/c
ID   AAI30331 standard; DNA; 31 BP.
XX
AC   AAI30331;
XX
DT   18-OCT-2001  (first entry)
XX
DE   Human single nucleotide polymorphism (SNP) RPS6KA1 5.
XX
KW   Human; resequence; genotype; disease; forensic; paternity testing;
KW   single nucleotide polymorphism; SNP; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(16,A)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200166800-A2.
XX
PD   13-SEP-2001.
XX
PF   07-MAR-2001; 2001WO-US07268.
XX
PR   07-MAR-2000; 2000US-0187510.
PR   22-MAY-2000; 2000US-0206129.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-522952/57.
XX
PT   Nucleic acid molecules from the human genome which include polymorphic
PT   sites, useful in methods for predicting the presence, absence or
PT   severity of a particular phenotype or disorder (e.g. diabetes)
PT   associated with a particular genotype -
XX
PS   Claim 1; Page 79; 145pp; English.
XX
CC   The invention relates to the identification of nucleic acid molecules
CC   (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC   which can predispose individuals to disease. Various genes from a number
CC   of individuals were resequenced and single nucleotide polymorphisms
CC   (SNPs) in these genes discovered. The method is useful for predicting the
CC   presence, absence or severity of a particular phenotype or disorder (e.g.
CC   diabetes) associated with a particular genotype. The nucleic acids
CC   containing the polymorphic sites may be useful in forensics and paternity
CC   testing.
XX
SQ   Sequence 31 BP; 5 A; 9 C; 14 G; 3 T; 0 other;

  Query Match             67.0%;  Score 13.4;  DB 22;  Length 31;
  Best Local Similarity   93.3%;  Pred. No. 3.1e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 CGCCCTGTGTCCCGA 20
           ||||||||| |||||
Db      28 CGCCCTGTGGCCCGA 14


RESULT 10
AAH44305
ID   AAH44305 standard; DNA; 33 BP.
XX
AC   AAH44305;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human fibrinogen 9 PCR primer SEQ ID NO:5.
XX
KW   Human; fibrinogen 9; cytostatic; haemostatic; antiinflammatory;
KW   immunomodulatory; virucide; malignant tumour; haemopathy;
KW   HIV infection; immunological disease; inflammation; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200148196-A1.
XX
PD   05-JUL-2001.
XX
PF   18-DEC-2000; 2000WO-CN00611.
XX
PR   24-DEC-1999;   99CN-0125762.
XX
PA   (BIOW-) BIOWINDOW GENE DEV LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2001-418278/44.
XX
```

```
PT   Fibrinogen 9 and encoded polynucleotide, applicable in diagnosis and
PT   treatment of malignant tumor, hemopathy, HIV infection, immunological
PT   diseases and various inflammations -
XX
PS   Example 5; Page 12; 27pp; Chinese.
XX
CC   The present sequence represents a PCR primer for human fibrinogen 9.
CC   Human fibrinogen 9 has cytostatic, haemostatic, antiinflammatory,
CC   immunomodulatory and virucide activities. The human fibrinogen 9 protein
CC   and polynucleotide sequences are applicable in the diagnosis and
CC   treatment of malignant tumour, haemopathy, HIV infection, immunological
CC   diseases and various inflammations.
XX
SQ   Sequence 33 BP; 6 A; 10 C; 10 G; 7 T; 0 other;

  Query Match             67.0%;  Score 13.4;  DB 22;  Length 33;
  Best Local Similarity   93.3%;  Pred. No. 3.1e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GGCGCCCTGTGTCCC 18
           || ||||||||||||
Db      18 GGAGCCCTGTGTCCC 32


RESULT 11
ABK36795
ID   ABK36795 standard; DNA; 90 BP.
XX
AC   ABK36795;
XX
DT   08-MAY-2002  (first entry)
XX
DE   Human DNA encoding MC1R segment 18.
XX
KW   Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW   viral infection; human immunodeficieny virus; melanoma;
KW   bacterial infection; Salmonella; Legionella; parasitic infection;
KW   Trypanosoma; Toxoplasma; Giardia; ds.
XX
OS   Homo sapiens.
XX
PN   WO200190197-A1.
XX
PD   29-NOV-2001.
XX
PF   25-MAY-2001; 2001WO-AU00622.
XX
PR   26-MAY-2000; 2000AU-0007761.
XX
PA   (AUSU ) UNIV AUSTRALIAN NAT.
XX
PI   Thomson SA,  Ramshaw IA;
XX
DR   WPI; 2002-147575/19.
DR   P-PSDB; AAU84975.
XX
PT   New synthetic polypeptides having several different segments of at
PT   least one parent polypeptide linked together differently compared to
PT   the linkage in the parent polypeptide, for inducing immune response
PT   against a pathogen or cancer  -
XX
PS   Example 3; Fig 27; 364pp; English.
XX
CC   The invention relates to a new synthetic polypeptide (I) comprising
CC   several different segments of at least one parent polypeptide linked
CC   together in a different relationship relative to their linkage in the
CC   parent polypeptide to impede, abrogate or otherwise alter at least one
CC   function associated with the parent polypeptide and for inducing an
CC   immune response against a pathogen or cancer. Also included are a
CC   synthetic polynucleotide encoding and a computer system for
CC   designing the synthetic polypeptides. The synthetic polypeptides and
CC   polynucleotides are referred to as a Savine. The synthetic polypeptide is
CC   useful for modulating immune responses preferably directed against a
CC   pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,
CC   colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone
CC   liver, oesophagus, brain, testicle, uterus), as potentiating agents.
CC   Compositions comprising the polypeptide may be used in the treatment or
CC   prophylaxis against viral (such as infections caused by HIV (human
CC   immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC   virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC   (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC   Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC   (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC   Trypanosoma, Toxoplasma and Giardia) infections. The present
CC   sequence encodes a peptide derived from a parent protein used to
CC   construct a savine of the invention.
XX
SQ   Sequence 90 BP; 15 A; 30 C; 14 G; 31 T; 0 other;

  Query Match             67.0%;  Score 13.4;  DB 24;  Length 90;
  Best Local Similarity   93.3%;  Pred. No. 3.2e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 CGCCCTGTGTCCCGA 20
           || ||||||||||||
Db      33 CGTCCTGTGTCCCGA 47


RESULT 12
ABL94381/c
ID   ABL94381 standard; DNA; 20 BP.
XX
AC   ABL94381;
XX
DT   29-JUL-2002  (first entry)
XX
DE   Mouse C/EBP beta phosphorothioate antisense oligonucleotide, SEQ ID:147.
XX
KW   Mouse; murine; C/EBP beta; CCAAT/enhancer-binding protein beta; C/EPB2;
KW   LAP; TCF5; CRP2; NFIL6; IL6DBP; NF-M; AGP/EBP; Apc/EBP;
KW   transcription factor; tissue development; cellular function;
KW   proliferation; differentiation; hormone responsiveness;
KW   oxidative stress response; IL-6 signalling mediator; interleukin-6;
KW   carbohydrate metabolism; immunity; Th1 response; female fertility;
KW   gluconeogenesis; ovarian; cancer; tumour formation; type II; diabetes;
KW   infection; inflammation; expression inhibition; phosphorothioate;
KW   antisense oligonucleotide; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1..20
FT                   /*tag=  a
FT                   /mod_base= OTHER
FT                   /note= "Phosphorothioate linkages"
FT   modified_base   1..5
FT                   /*tag=  b
FT                   /mod_base= OTHER
FT                   /note= "2'-methoxyethyl (2'-MOE) nucleotides. All 2' MOE
FT                          cytosines are 5-methylcytosine"
FT   modified_base   16..20
FT                   /*tag=  c
FT                   /mod_base= OTHER
FT                   /note= "2'-methoxyethyl (2'-MOE) nucleotides. All 2' MOE
FT                          cytosines are 5-methylcytosine"
XX
PN   US6271030-B1.
XX
PD   07-AUG-2001.
XX
PF   14-JUN-2000; 2000US-0593711.
XX
PR   14-JUN-2000; 2000US-0593711.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
```

```
PI   Monia BP,  Butler MM,  Wyatt J;
XX
DR   WPI; 2002-214451/27.
XX
PT   Novel antisense compound targeted to nucleic acids encoding human or
PT   mouse CCAAT/enhancer binding protein (C/EBP) beta, useful in vitro for
PT   inhibiting expression of human or mouse C/EBP beta in cells/tissues  -
XX
PS   Claim 1; Column 49-50; 69pp; English.
XX
CC   Sequences ABL94252-ABL94476 represent antisense oligonucleotides targeted
CC   to the human or mouse CCAAT/enhancer-binding protein alpha (C/EBP alpha)
CC   gene, which inhibit its expression. The antisense oligonucleotides were
CC   designed to target different regions of the human and/or mouse C/EBP
CC   alpha RNA, and were analysed for their effect on C/EBP alpha mRNA levels
CC   by quantitative real-time PCR. The C/EBP family of proteins are a family
CC   of transcription factors which regulate the expression of a wide range
CC   of genes that control normal tissue development, cellular function,
CC   cellular proliferation and functional differentiation. C/EBP beta (also
CC   known as C/EPB2, LAP, TCF5, CRP2, NFIL6, IL6DBP, NF-M, AGP/EBP and
CC   Apc/EBP) primarily regulates hormone responsiveness and oxidative stress
CC   responses and is a mediator of IL-6 (interleukin-6) signalling. C/EBP
CC   beta is thought to be involved in carbohydrate metabolism, immunity, the
CC   Th1 response, female fertility and gluconeogenic pathways. C/EBP beta is
CC   expressed in the liver, lung, spleen, kidney, brain, and testis, with the
CC   highest expression found in the lung. It is also expressed at a higher
CC   level in malignant ovarian tissue compared with normal ovarian tissue,
CC   and its expression in pancreas is upregulated in response to chronically
CC   elevated levels of glucose, indicating that it is involved in the
CC   impairment of insulin secretion in type II diabetes. The oligonucleotides
CC   of the invention are useful for diagnosis, prevention and treatment of
CC   conditions associated with C/EBP beta expression, such as cancer
CC   (particularly ovarian cancer), tumour formation, diabetes (particularly
CC   type II diabetes), infection, or inflammation.
XX
SQ   Sequence 20 BP; 6 A; 7 C; 6 G; 1 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 24;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 3.8e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATGGCGCCCTGTGTCCCG 19
           |||||| || || |||||
Db      20 ATGGCGGCCGGTTTCCCG 3


RESULT 13
AAT90699
ID   AAT90699 standard; DNA; 34 BP.
XX
AC   AAT90699;
XX
DT   05-JAN-1998  (first entry)
XX
DE   Primer #4 for env containing plasmid construction.
XX
KW   Packaging-deficient construct; viral gag-pol gene; packaging cell line;
KW   moloney murine leukaemia virus; MoMLV; viral env gene; helper construct;
KW   gene therapy; polymerase chain reaction; primer; amplify; PCR; ss.
XX
OS   Synthetic.
XX
PN   WO9708330-A1.
XX
PD   06-MAR-1997.
XX
PF   23-AUG-1996;   96WO-GB02061.
XX
PR   23-AUG-1995;   95GB-0017263.
XX
PA   (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
PI   Collins MKL,  Cosset F,  Takeuchi Y,  Weiss RA;
XX
DR   WPI; 1997-179287/16.
XX
PT   Selectable retroviral packaging cell lines and expression constructs
PT   - comprise selectable gene downstream of gene of interest, are
PT   selectable due to the in-efficiency associated with translation
PT   re-initiation
XX
PS   Example 2; Page 19; 79pp; English.
XX
CC   AAT90696-T90701 represent amplification primers used in the construction
CC   of viral env gene containing recombinant expression vectors (REV) of the
CC   invention. The vectors are used to create a packaging cell line. The
CC   REVs of the invention comprise a gene of interest (GOI) and a SM gene.
CC   The SM gene is arranged downstream of the GOI and a GOI associated stop
CC   codon is spaced from a start codon of the SM gene to ensure that the SM
CC   protein is expressed as a result of translation reinitiation. The cell
CC   lines are transformed with two REVs, both are replication deficient, one
CC   contains the viral gag-pol gene, the other the viral env gene. By using
CC   helper constructs, such as the REV's, which are directly selectable and
CC   which provide for high expression of the viral gene, high titre
CC   retroviral vectors may be obtained. The packaging cell lines are useful
CC   for gene therapy. Prior packaging cell lines using full length retroviral
CC   genomes as helper genomes were isolated by cotransfecting them with
CC   plasmids encoding selectable markers. However, the helper functions can
CC   be lost during the passages of the cells in culture and the current
CC   packaging systems provide limited titres of infectious retroviral
CC   vectors. Co-transfection with a plasmid encoding a SM does not directly
CC   select the best gag-pol-env-expressing cells. The new retroviral
CC   packaging cell lines overcome these problems.
XX
SQ   Sequence 34 BP; 6 A; 10 C; 9 G; 9 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 18;  Length 34;
  Best Local Similarity   83.3%;  Pred. No. 3.9e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATGGCGCCCTGTGTCCCG 19
           |||||||||| || | ||
Db       5 ATGGCGCCCTATGGCTCG 22


RESULT 14
AAD14556/c
ID   AAD14556 standard; DNA; 40 BP.
XX
AC   AAD14556;
XX
DT   01-NOV-2001  (first entry)
XX
DE   Arabidopsis thaliana ABF binding sequence #3.
XX
KW   Abscisic acid responsive element; ABRE; ABRE-binding factor; ABF;
KW   stress treatment; transgenic plant; environmental stress; ds.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    12..22
FT                   /*tag= a
FT                   /note= "Conserved region"
XX
PN   US6232461-B1.
XX
PD   15-MAY-2001.
XX
PF   19-SEP-2000; 2000US-0665309.
XX
PR   12-OCT-1999;   99US-0416050.
XX
PA   (KOKU-) KOREA KUMHO PETROCHEMICAL CO LTD.
XX
PI   Kim SY;
```

```
XX
DR   WPI; 2001-366358/38.
XX
PT   New nucleic acid encoding abscisic acid responsive element binding
PT   factor 4, useful for generating transgenic plants, which are tolerant
PT   to multiple environmental stresses -
XX
PS   Disclosure; Fig 5B; 42pp; English.
XX
CC   The invention relates to nucleic acid encoding the Abscisic acid
CC   responsive element (ABRE)-binding factor 4 (ABF4). ABF4 belongs to
CC   the ABF family of factors which bind abscisic acid responsive elements
CC   in plants. Expression of ABFs is inducible by abscisic acid and various
CC   stress treatments. ABFs have the potential to activate a large number
CC   of abscisic acid/stress responsive genes and thus a nucleic acid
CC   molecule encoding ABF4 is useful for generating transgenic plants
CC   that are tolerant to multiple environmental stresses. The present
CC   sequence is Arabidopsis thaliana ABF binding sequence which contains
CC   the group IA consensus sequence.
XX
SQ   Sequence 40 BP; 9 A; 11 C; 14 G; 6 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 22;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 3.9e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGGCGCCCTGTGTCCCGA 20
           | |||||  |||||||||
Db      26 TTGCGCCACGTGTCCCGA 9


RESULT 15
AAS07861/c
ID   AAS07861 standard; DNA; 40 BP.
XX
AC   AAS07861;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Binding site selection assay group IA sequence #3.
XX
KW   Abscisic acid responsive element binding factor; ABF; bZIP; ss; G/ABRE;
KW   G-box-like ABRE; coupling element-like ABRE; C/ABRE; transgenic plant;
KW   stress responsive gene; environmental stress; cress;
KW   binding site selection assay.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    13..22
FT                   /*tag=  a
FT                   /note= "Group IA consensus sequence"
XX
PN   EP1097999-A2.
XX
PD   09-MAY-2001.
XX
PF   29-JUN-2000; 2000EP-0305459.
XX
PR   04-NOV-1999;   99KR-0048477.
XX
PA   (KOKU-) KOREA KUMHO PETROCHEMICAL CO LTD.
XX
PI   Kim SY;
XX
DR   WPI; 2001-337113/36.
XX
PT   Nucleic acid encoding novel transcription factors that bind to various
PT   abscisic acid responsive elements, for generating stress tolerant
PT   transgenic plants  -
XX
PS   Disclosure; Fig 5B; 37pp; English.
XX
CC   The sequence represents an oligonucleotide studied in a binding site
CC   selection assay, for studying abscisic acid responsive element binding
CC   factor (ABF) proteins. The oligonucleotide contains the IA consensus
CC   sequence. ABFs are bZIP class transcription factors that can bind to two
CC   classes of ABRE, namely G-box-like ABREs (G/ABRE) and coupling
CC   element-like ABREs (C/ABRE). Expression of ABF is inducible by abscisic
CC   acid and various stress treatments and they can transactivate an
CC   ABRE-containing reporter gene in yeast. Therefore, the ABFs are useful
CC   for activating a large number of abscisic acid or stress responsive genes
CC   and for generating transgenic plants that are tolerant to multiple
CC   environmental stresses.
XX
SQ   Sequence 40 BP; 9 A; 11 C; 14 G; 6 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 22;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 3.9e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGGCGCCCTGTGTCCCGA 20
           | |||||  |||||||||
Db      26 TTGCGCCACGTGTCCCGA 9


Search completed: November 23, 2002, 06:28:48
Job time : 102.6 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-7
Perfect score: 20
Sequence: 1 aatggcgccctgtgtcccga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.2	71.0	25	1	US-08-182-961B-10	Sequence 10, Appl
	2	14.2	71.0	25	4	US-09-007-678B-10	Sequence 10, Appl
c	3	13.2	66.0	20	4	US-09-593-711A-147	Sequence 147, App
	4	13.2	66.0	34	4	US-09-011-745-13	Sequence 13, Appl
c	5	13.2	66.0	40	4	US-09-416-050A-35	Sequence 35, Appl
c	6	13.2	66.0	40	4	US-09-664-800-35	Sequence 35, Appl
c	7	13.2	66.0	40	4	US-09-665-309-35	Sequence 35, Appl
c	8	13.2	66.0	40	4	US-09-661-569-35	Sequence 35, Appl
c	9	12.8	64.0	21	1	US-08-009-263C-9	Sequence 9, Appli
	10	12.8	64.0	21	1	US-08-009-263C-72	Sequence 72, Appl
c	11	12.8	64.0	21	1	US-07-927-506-9	Sequence 9, Appli
c	12	12.8	64.0	21	3	US-08-838-715B-9	Sequence 9, Appli
	13	12.8	64.0	21	3	US-08-838-715B-72	Sequence 72, Appl
c	14	12.8	64.0	21	5	PCT-US91-05815-9	Sequence 9, Appli
c	15	12.6	63.0	65	2	US-08-418-848A-70	Sequence 70, Appl
	16	12.6	63.0	72	2	US-08-749-852-48	Sequence 48, Appl
	17	12.6	63.0	73	2	US-08-749-852-50	Sequence 50, Appl
c	18	12.2	61.0	20	3	US-09-428-584-13	Sequence 13, Appl
	19	12.2	61.0	86	1	US-07-964-624D-82	Sequence 82, Appl
	20	12.2	61.0	86	1	US-08-442-062-82	Sequence 82, Appl
	21	12.2	61.0	86	1	US-08-748-697A-82	Sequence 82, Appl
	22	12.2	61.0	86	4	US-09-165-616-82	Sequence 82, Appl
c	23	12	60.0	40	4	US-09-416-050A-37	Sequence 37, Appl
c	24	12	60.0	40	4	US-09-664-800-37	Sequence 37, Appl
c	25	12	60.0	40	4	US-09-665-309-37	Sequence 37, Appl
c	26	12	60.0	40	4	US-09-661-569-37	Sequence 37, Appl
c	27	12	60.0	45	3	US-08-612-973-98	Sequence 98, Appl
c	28	12	60.0	45	4	US-08-927-597-98	Sequence 98, Appl
c	29	12	60.0	58	5	PCT-US95-02917-3	Sequence 3, Appli
c	30	12	60.0	62	2	US-08-850-049-14	Sequence 14, Appl
c	31	12	60.0	62	2	US-08-050-478-14	Sequence 14, Appl
c	32	12	60.0	62	4	US-07-858-747B-14	Sequence 14, Appl
c	33	12	60.0	62	4	US-09-414-117-14	Sequence 14, Appl
c	34	12	60.0	62	4	US-09-678-437-14	Sequence 14, Appl
	35	12	60.0	70	1	US-08-400-440A-82	Sequence 82, Appl
	36	12	60.0	70	1	US-08-463-093A-82	Sequence 82, Appl
	37	12	60.0	70	2	US-08-460-888A-82	Sequence 82, Appl
	38	12	60.0	70	2	US-08-894-578-82	Sequence 82, Appl
	39	12	60.0	70	4	US-09-412-017-82	Sequence 82, Appl
	40	12	60.0	71	2	US-08-465-591A-39	Sequence 39, Appl
	41	12	60.0	71	2	US-08-465-594A-39	Sequence 39, Appl
	42	12	60.0	71	4	US-08-973-124-224	Sequence 224, App
	43	12	60.0	71	5	PCT-US96-08014-224	Sequence 224, App
c	44	11.8	59.0	27	4	US-09-253-396A-134	Sequence 134, App
	45	11.8	59.0	28	4	US-09-552-322-60	Sequence 60, Appl

ALIGNMENTS

```
RESULT 1
US-08-182-961B-10
; Sequence 10, Application US/08182961B
; Patent No. 5677125
;  GENERAL INFORMATION:
;    APPLICANT:  HOLT, JEFFREY T.
;    APPLICANT:  JENSEN, ROY A.
;    APPLICANT:  PAGE, DAVID L.
;    APPLICANT:  OBERMILLER, PATRICE S.
;    APPLICANT:  ROBINSON-BENION, CHERYL L.
;    TITLE OF INVENTION:  METHOD OF DETECTION AND DIAGNOSIS OF PRE-INVASIVE CANC
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  I.C. WADDEY, JR.
;      STREET:  27TH FLOOR, L & C TOWER, 401 CHURCH
;      CITY:  NASHVILLE
;      STATE:  TENNESSE
;      COUNTRY:  USA
;      ZIP:  37219
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 800 kB storage
;      COMPUTER:  IBM PC/XT/AT compatible
;      OPERATING SYSTEM:  MS-DOS (version 5.0)
;      SOFTWARE:  WordPerfect 5.1/WordPerfect Editor
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/182,961B
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  I.C. WADDEY, JR.
;      REGISTRATION NUMBER:  25,180
;      REFERENCE/DOCKET NUMBER:  0216-9409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (615) 242-2400
;      TELEFAX:  (615) 242-2221
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;      DESCRIPTION:  PCR primer
;    HYPOTHETICAL:  yes
;    ANTI-SENSE:  no
;    FRAGMENT TYPE:  oligonucleotide
```

US-08-182-961B-10

Query Match 71.0%; Score 14.2; DB 1; Length 25;
Best Local Similarity 84.2%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      2 ATGGCGCCCTGTGTCCCGA 20
          |||||| |||||  |||||
Db      3 ATGGCGTCCTGTAACCCGA 21
```

RESULT 2
US-09-007-678B-10
; Sequence 10, Application US/09007678B
; Patent No. 6342483
; GENERAL INFORMATION:
; APPLICANT: HOLT, JEFFREY T.
; APPLICANT: JENSEN, ROY A.
; APPLICANT: PAGE, DAVID L.
; APPLICANT: OBERMILLER, PATRICE S.
; APPLICANT: ROBINSON-BENION, CHERYL L.
; APPLICANT: THOMPSON, MARILYN E.
; TITLE OF INVENTION: METHOD FOR DETECTION AND TREATMENT OF BREAST CANCER
; FILE REFERENCE: Attorney Docket No. 6342483 1242-1-2-2
; CURRENT APPLICATION NUMBER: US/09/007,678B
; CURRENT FILING DATE: 1998-01-15
; PRIOR APPLICATION NUMBER: 08/373,799
; PRIOR FILING DATE: 1995-01-17
; PRIOR APPLICATION NUMBER: 08/182,961
; PRIOR FILING DATE: 1994-01-14
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: MicroSoft WordPad
; SEQ ID NO 10
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized PCR Primer
US-09-007-678B-10

Query Match 71.0%; Score 14.2; DB 4; Length 25;
Best Local Similarity 84.2%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      2 ATGGCGCCCTGTGTCCCGA 20
          |||||| |||||  |||||
Db      3 ATGGCGTCCTGTAACCCGA 21
```

RESULT 3
US-09-593-711A-147/c
; Sequence 147, Application US/09593711A
; Patent No. 6271030
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Madeline M. Butler
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF C/EBP BETA EXPRESSION
; FILE REFERENCE: RTS-0118
; CURRENT APPLICATION NUMBER: US/09/593,711A
; CURRENT FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 244
; SEQ ID NO 147
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-593-711A-147

Query Match 66.0%; Score 13.2; DB 4; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.6e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      2 ATGGCGCCCTGTGTCCCG 19
          |||||| || || |||||
Db     20 ATGGCGGCCGGTTTCCCG 3
```

RESULT 4
US-09-011-745-13
; Sequence 13, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; CURRENT FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 34
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Oligonucleotide
US-09-011-745-13

Query Match 66.0%; Score 13.2; DB 4; Length 34;
Best Local Similarity 83.3%; Pred. No. 6.7e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      2 ATGGCGCCCTGTGTCCCG 19
          |||||||||| || | ||
Db      5 ATGGCGCCCTATGGCTCG 22
```

RESULT 5
US-09-416-050A-35/c
; Sequence 35, Application US/09416050A
; Patent No. 6194559
; GENERAL INFORMATION:
; APPLICANT: KIM. Soo Young
; TITLE OF INVENTION: Abscisic Acid Responsive Element -Binding Transcription Facto
; FILE REFERENCE: 1942/42
; CURRENT APPLICATION NUMBER: US/09/416,050A
; CURRENT FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-416-050A-35

Query Match 66.0%; Score 13.2; DB 4; Length 40;
Best Local Similarity 83.3%; Pred. No. 6.7e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      3 TGGCGCCCTGTGTCCCGA 20
          | |||||  |||||||||
Db     26 TTGCGCCACGTGTCCCGA 9
```

RESULT 6
US-09-664-800-35/c
; Sequence 35, Application US/09664800

```
; Patent No. 6218527
; GENERAL INFORMATION:
;  APPLICANT:    KIM. Soo Young
;  TITLE OF INVENTION:    Abscisic Acid Responsive Element -Binding Transcription Factor
;  FILE REFERENCE:   1942/42
;  CURRENT APPLICATION NUMBER: US/09/664,800
;  CURRENT FILING DATE:  2000-09-19
;  PRIOR APPLICATION NUMBER:   09/416,050
;  PRIOR FILING DATE:   1999-10-12
;  NUMBER OF SEQ ID NOS:    83
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 35
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-664-800-35

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 6.7e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 TGGCGCCCTGTGTCCCGA 20
          | |||||  |||||||||
Db     26 TTGCGCCACGTGTCCCGA 9


RESULT 7
US-09-665-309-35/c
; Sequence 35, Application US/09665309
; Patent No. 6232461
; GENERAL INFORMATION:
;  APPLICANT:    KIM. Soo Young
;  TITLE OF INVENTION:    Abscisic Acid Responsive Element -Binding Transcription Factor
;  FILE REFERENCE:   1942/42
;  CURRENT APPLICATION NUMBER: US/09/665,309
;  CURRENT FILING DATE:  2000-09-19
;  PRIOR APPLICATION NUMBER:   09/416,050
;  PRIOR FILING DATE:   1999-10-12
;  NUMBER OF SEQ ID NOS:    83
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 35
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-665-309-35

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 6.7e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 TGGCGCCCTGTGTCCCGA 20
          | |||||  |||||||||
Db     26 TTGCGCCACGTGTCCCGA 9


RESULT 8
US-09-661-569-35/c
; Sequence 35, Application US/09661569
; Patent No. 6245905
; GENERAL INFORMATION:
;  APPLICANT:    KIM. Soo Young
;  TITLE OF INVENTION:    Abscisic Acid Responsive Element -Binding Transcription Factor
;  FILE REFERENCE:   1942/42
;  CURRENT APPLICATION NUMBER: US/09/661,569
;  CURRENT FILING DATE:  2000-09-14
;  PRIOR APPLICATION NUMBER:   09/416,050
;  PRIOR FILING DATE:   1999-10-12
;  NUMBER OF SEQ ID NOS:    83
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 35
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-661-569-35

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 40;
  Best Local Similarity   83.3%;  Pred. No. 6.7e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 TGGCGCCCTGTGTCCCGA 20
          | |||||  |||||||||
Db     26 TTGCGCCACGTGTCCCGA 9


RESULT 9
US-08-009-263C-9/c
; Sequence 9, Application US/08009263C
; Patent No. 5442049
;  GENERAL INFORMATION:
;    APPLICANT:  Kevin Anderson, Kenneth Draper, Brenda Baker
;    TITLE OF INVENTION:  Oligonucleotides for Modulating the
;    TITLE OF INVENTION:  Effects of Cytomegalovirus Infections
;    NUMBER OF SEQUENCES:  88
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5442049ris
;      STREET:  One Liberty Place -- 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/009,263C
;      FILING DATE:  January 25, 1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  927,506
;      FILING DATE:  No. 5442049ember 19, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISIS-0844
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
US-08-009-263C-9

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 GCGCCCTGTGTCCCGA 20
          |||||| |||||| ||
Db     21 GCGCCCGGTGTCCGGA 6


RESULT 10
US-08-009-263C-72
; Sequence 72, Application US/08009263C
; Patent No. 5442049
;  GENERAL INFORMATION:
;    APPLICANT:  Kevin Anderson, Kenneth Draper, Brenda Baker
```

```
;    TITLE OF INVENTION:  Oligonucleotides for Modulating the
;    TITLE OF INVENTION:  Effects of Cytomegalovirus Infections
;    NUMBER OF SEQUENCES:  88
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5442049ris
;      STREET:  One Liberty Place -- 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/009,263C
;      FILING DATE:  January 25, 1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  927,506
;      FILING DATE:  No. 5442049ember 19, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISIS-0844
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-009-263C-72

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GCGCCCTGTGTCCCGA 20
            |||||| |||||| ||
Db       1 GCGCCCGGTGTCCGGA 16


RESULT 11
US-07-927-506-9/c
; Sequence 9, Application US/07927506
; Patent No. 5591720
;  GENERAL INFORMATION:
;    APPLICANT:  Anderson, Kevin P.
;    APPLICANT:  Draper, Kenneth G.
;    TITLE OF INVENTION:  Oligonucleotides for Modulating
;    TITLE OF INVENTION:  the Effects of Cytomegalovirus Infections
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  No. 5591720ris
;      STREET:  One Liberty Place -- 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/927,506
;      FILING DATE:  19921119
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Licata, Jane M.
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISIS-0408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
US-07-927-506-9

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GCGCCCTGTGTCCCGA 20
            |||||| |||||| ||
Db      21 GCGCCCGGTGTCCGGA 6


RESULT 12
US-08-838-715B-9/c
; Sequence 9, Application US/08838715B
; Patent No. 6153595
;  GENERAL INFORMATION:
;    APPLICANT:  Draper, Chapman, Kisner, Anderson
;    TITLE OF INVENTION:  Composition and Method for Treatment
;    TITLE OF INVENTION:  of CMV Infection
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  66 E. Main Street
;      CITY:  Marlton
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/838,715B
;      FILING DATE:  April 9, 1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/568,366
;      FILING DATE:  8/16/90
;      APPLICATION NUMBER:  07/927,506
;      FILING DATE:  11/19/92
;      APPLICATION NUMBER:  08/009,263
;      FILING DATE:  1/25/93
;      APPLICATION NUMBER:  08/233,711
;      FILING DATE:  4/26/94
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISPH-0204
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 810-1454
```

```
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
US-08-838-715B-9

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GCGCCCTGTGTCCCGA 20
           |||||| |||||| ||
Db      21 GCGCCCGGTGTCCGGA 6


RESULT 13
US-08-838-715B-72
; Sequence 72, Application US/08838715B
; Patent No. 6153595
;  GENERAL INFORMATION:
;    APPLICANT:  Draper, Chapman, Kisner, Anderson
;    TITLE OF INVENTION:  Composition and Method for Treatment
;    TITLE OF INVENTION:  of CMV Infection
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  66 E. Main Street
;      CITY:  Marlton
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/838,715B
;      FILING DATE:  April 9, 1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/568,366
;      FILING DATE:  8/16/90
;      APPLICATION NUMBER:  07/927,506
;      FILING DATE:  11/19/92
;      APPLICATION NUMBER:  08/009,263
;      FILING DATE:  1/25/93
;      APPLICATION NUMBER:  08/233,711
;      FILING DATE:  4/26/94
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISPH-0204
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 810-1454
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-838-715B-72

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GCGCCCTGTGTCCCGA 20
           |||||| |||||| ||
Db       1 GCGCCCGGTGTCCGGA 16


RESULT 14
PCT-US91-05815-9/c
; Sequence 9, Application PC/TUS9105815
;  GENERAL INFORMATION:
;    APPLICANT:  Anderson, Kevin P.
;    APPLICANT:  Draper, Kenneth G.
;    TITLE OF INVENTION:  Oligonucleotides for Modulating
;    TITLE OF INVENTION:  the Effects of Cytomegalovirus Infections
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  Norris
;      STREET:  One Liberty Place -- 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb
;      MEDIUM TYPE:  STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/05815
;      FILING DATE:  19910814
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Licata, Jane M.
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISIS-0408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
PCT-US91-05815-9

  Query Match             64.0%;  Score 12.8;  DB 5;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GCGCCCTGTGTCCCGA 20
           |||||| |||||| ||
Db      21 GCGCCCGGTGTCCGGA 6


RESULT 15
US-08-418-848A-70/c
; Sequence 70, Application US/08418848A
; Patent No. 5847096
;  GENERAL INFORMATION:
;    APPLICANT:  SCHUBERT, MANFRED, HARMISON II,
;    APPLICANT:  GEORGE G., CHANG-JIE, CHEN, BANJERJEA, AKHIL
;    TITLE OF INVENTION:  DEFECTIVE, INTERFERING
;    TITLE OF INVENTION:  HIV PARTICLES
;    NUMBER OF SEQUENCES:  77
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       COUNTRY:  U.S.A.
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WORD PERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/418,848A
;       FILING DATE:  07-APR-1995
;       CLASSIFICATION:  526
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/936,849
;       FILING DATE:  28-AUG-1992
;       CLASSIFICATION:  526
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  RICHARD W. BORK
;       REGISTRATION NUMBER:  36,459
;       REFERENCE/DOCKET NUMBER:  2026-4091US2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-758-4800
;       TELEFAX:  212-751-6849
;       TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  70:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  65 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-418-848A-70

  Query Match             63.0%;  Score 12.6;  DB 2;  Length 65;
  Best Local Similarity   78.9%;  Pred. No. 1.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AATGGCGCCCTGTGTCCCG 19
           |||||| || |||||  ||
Db      56 AATGGCACCATGTGTTTCG 38


Search completed: November 23, 2002, 06:36:06
Job time : 22.55 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-7
Perfect score: 20
Sequence: 1 aatggcgccctgtgtcccga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	13	65.0	31	10	US-09-801-274-819	Sequence 819, App
c	2	12.2	61.0	41	10	US-09-778-168-17	Sequence 17, Appl
c	3	12.2	61.0	41	10	US-09-778-175-17	Sequence 17, Appl
c	4	12.2	61.0	41	10	US-09-335-218-17	Sequence 17, Appl
	5	11.8	59.0	30	12	US-10-028-051-17	Sequence 17, Appl
c	6	11.8	59.0	31	10	US-09-801-274-867	Sequence 867, App
c	7	11.8	59.0	52	10	US-09-146-157-2	Sequence 2, Appli
	8	11.8	59.0	72	10	US-09-242-980-32	Sequence 32, Appl
c	9	11.6	58.0	87	10	US-09-873-075A-5	Sequence 5, Appli
c	10	11.4	57.0	20	10	US-09-780-172-54	Sequence 54, Appl
c	11	11.4	57.0	28	10	US-09-987-967-3	Sequence 3, Appli
c	12	11.4	57.0	52	9	US-10-046-935-1880	Sequence 1880, Ap
	13	11.4	57.0	87	10	US-09-864-761-32231	Sequence 32231, A
	14	11.2	56.0	16	10	US-09-420-433-34	Sequence 34, Appl
	15	11.2	56.0	25	12	US-10-007-805-554	Sequence 554, App
c	16	11.2	56.0	28	9	US-09-905-291A-24	Sequence 24, Appl
c	17	11.2	56.0	28	10	US-09-909-320-24	Sequence 24, Appl
c	18	11.2	56.0	28	10	US-09-909-088B-24	Sequence 24, Appl
	19	11.2	56.0	34	10	US-09-801-368-13	Sequence 13, Appl
c	20	11.2	56.0	75	10	US-09-864-761-20571	Sequence 20571, A
	21	11.2	56.0	76	10	US-09-783-590-3546	Sequence 3546, Ap
c	22	11.2	56.0	85	10	US-09-864-761-24620	Sequence 24620, A
	23	11.2	56.0	89	9	US-09-933-797-225	Sequence 225, App
c	24	11.2	56.0	91	10	US-09-864-761-21111	Sequence 21111, A
c	25	11.2	56.0	94	10	US-09-864-761-31201	Sequence 31201, A
	26	11.2	56.0	96	10	US-09-864-761-29606	Sequence 29606, A
	27	11.2	56.0	97	10	US-09-864-761-19857	Sequence 19857, A
	28	11	55.0	24	10	US-09-965-703-65	Sequence 65, Appl
c	29	11	55.0	32	10	US-09-896-852-39	Sequence 39, Appl
c	30	11	55.0	36	9	US-09-251-330-6	Sequence 6, Appli
c	31	11	55.0	42	10	US-09-795-006A-5	Sequence 5, Appli
	32	11	55.0	42	10	US-09-795-006A-14	Sequence 14, Appl
	33	11	55.0	51	9	US-10-027-806-105	Sequence 105, App
	34	11	55.0	64	10	US-09-922-217-1029	Sequence 1029, Ap
	35	11	55.0	64	10	US-09-833-263-1029	Sequence 1029, Ap
	36	11	55.0	96	10	US-09-815-343-1009	Sequence 1009, Ap
	37	11	55.0	96	10	US-09-815-343-1137	Sequence 1137, Ap
c	38	11	55.0	98	10	US-09-864-761-26075	Sequence 26075, A
c	39	10.8	54.0	24	10	US-09-853-830-119	Sequence 119, App
	40	10.8	54.0	30	10	US-09-986-191-5	Sequence 5, Appli
	41	10.8	54.0	30	10	US-09-986-191-7	Sequence 7, Appli
	42	10.8	54.0	31	10	US-09-801-274-88	Sequence 88, Appl
c	43	10.8	54.0	31	10	US-09-801-274-893	Sequence 893, App
	44	10.8	54.0	37	10	US-09-250-611-106	Sequence 106, App
	45	10.8	54.0	40	9	US-09-826-025-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-801-274-819/c
; Sequence 819, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 819
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-819

Query Match 65.0%; Score 13; DB 10; Length 31;
Best Local Similarity 86.7%; Pred. No. 6.4e+02;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       6 CGCCCTGTGTCCCGA 20
           ||||||||| ||:||
Db      28 CGCCCTGTGGCCYGA 14
```

RESULT 2
US-09-778-168-17/c
; Sequence 17, Application US/09778168
; Patent No. US20010009761A1
; GENERAL INFORMATION:
; APPLICANT: Wright, David J.
; APPLICANT: Milla, Maria A.
; APPLICANT: Nadeau, James G.
; APPLICANT: Walker, G. Terrance

```
;  TITLE OF INVENTION: Methods for Detecting Nucleic Acid Sequence Variations
; Patent No. US20010009761A1
;  FILE REFERENCE: SNP Detection
;  CURRENT APPLICATION NUMBER: US/09/778,168
;  CURRENT FILING DATE:  2001-02-07
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 41
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-778-168-17

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 41;
  Best Local Similarity   82.4%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 GGCGCCCTGTGTCCCGA 20
          | |||| | ||||||||
Db     35 GTCGCCGTTTGTCCCGA 19


RESULT 3
US-09-778-175-17/c
; Sequence 17, Application US/09778175
; Patent No. US20010039334A1
; GENERAL INFORMATION:
;  APPLICANT: Wright, David J.
;  APPLICANT:  Milla, Maria A.
;  APPLICANT:  Nadeau, James G.
;  APPLICANT:  Walker, G. Terrance
;  TITLE OF INVENTION: Methods for Detecting Nucleic Acid Sequence Variations
;  FILE REFERENCE: SNP Detection
;  CURRENT APPLICATION NUMBER: US/09/778,175
;  CURRENT FILING DATE:  2001-02-07
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 41
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-778-175-17

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 41;
  Best Local Similarity   82.4%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 GGCGCCCTGTGTCCCGA 20
          | |||| | ||||||||
Db     35 GTCGCCGTTTGTCCCGA 19


RESULT 4
US-09-335-218-17/c
; Sequence 17, Application US/09335218
; Patent No. US20020025519A1
; GENERAL INFORMATION:
;  APPLICANT: Wright, David J.
;  APPLICANT: Milla, Maria A.
;  APPLICANT: Nadeau, James G.
;  APPLICANT: Walker, G. Terrance
;  TITLE OF INVENTION: Methods for Detecting Nucleic Acid Sequence Variations
;  FILE REFERENCE: SNP Detection
;  CURRENT APPLICATION NUMBER: US/09/335,218
;  CURRENT FILING DATE: 1999-06-17
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 41
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-335-218-17

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 41;
  Best Local Similarity   82.4%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 GGCGCCCTGTGTCCCGA 20
          | |||| | ||||||||
Db     35 GTCGCCGTTTGTCCCGA 19


RESULT 5
US-10-028-051-17
; Sequence 17, Application US/10028051
; Patent No. US20020147329A1
;   GENERAL INFORMATION:
;        APPLICANT: Luyten, Frank P.
;                   Hoang, Bang
;                   Moos, Jr., Malcolm
;                   Wang, Shouwen
;        TITLE OF INVENTION: ISOLATION AND USE OF TISSUE
;                            GROWTH INDUCING FRZB PROTEIN
;        NUMBER OF SEQUENCES: 23
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Knobbe, Martens, Olson & Bear
;             STREET: 620 Newport Center Drive, 16th Floor
;             CITY: Newport Beach
;             STATE: CA
;             COUNTRY: U.S.A.
;             ZIP: 92660
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ Version 1.5
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/028,051
;             FILING DATE: 19-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/729,452
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Bartfeld, Neil S
;             REGISTRATION NUMBER: 39,901
;             REFERENCE/DOCKET NUMBER: NIH133.001A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 619-235-8550
;             TELEFAX: 619-235-0176
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 17:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 30 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-028-051-17

  Query Match             59.0%;  Score 11.8;  DB 12;  Length 30;
  Best Local Similarity   86.7%;  Pred. No. 2.5e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 TGGCGCCCTGTGTCC 17
          ||||  |||||||||
Db      5 TGGCTGCCTGTGTCC 19


RESULT 6
US-09-801-274-867/c
; Sequence 867, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
```

```
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 867
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-867

  Query Match            59.0%;  Score 11.8;  DB 10;  Length 31;
  Best Local Similarity  86.7%;  Pred. No. 2.5e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GGCGCCCTGTGTCCC 18
           || |||||||||| |
Db      31 GGAGCCCTGTGTCTC 17


RESULT 7
US-09-146-157-2/c
; Sequence 2, Application US/09146157
; Patent No. US20010009760A1
; GENERAL INFORMATION:
;  APPLICANT: HORN, Thomas
;  APPLICANT: SCHROEDER, Hartmut R.
;  APPLICANT: WARNER, Brian D.
;  APPLICANT: FISS, Ellen
;  APPLICANT: SELLS, Todd
;  APPLICANT: LAW, Say-Jong
;  TITLE OF INVENTION: OLIGONUCLEOTIDE PROBES BEARING QUENCHABLE FLUORESCENT LABELS,
;  TITLE OF INVENTION: AND METHODS OF USE THEREOF
;  FILE REFERENCE: 1411.002
;  CURRENT APPLICATION NUMBER: US/09/146,157
;  CURRENT FILING DATE: 1998-09-03
;  EARLIER APPLICATION NUMBER: 60/057,810
;  EARLIER FILING DATE: 1997-09-04
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 52
;   TYPE: DNA
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism:This information
;   OTHER INFORMATION: is not available.
US-09-146-157-2

  Query Match            59.0%;  Score 11.8;  DB 10;  Length 52;
  Best Local Similarity  86.7%;  Pred. No. 2.6e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GGCGCCCTGTGTCCC 18
           |||| ||||| ||||
Db      39 GGCGTCCTGTATCCC 25


RESULT 8
US-09-242-980-32
; Sequence 32, Application US/09242980
; Patent No. US20010021073A1
; GENERAL INFORMATION:
;  APPLICANT:  Ogi, Kasuhiro
;  APPLICANT:  Onda, Haruo
;  TITLE OF INVENTION:  NOVEL PROTEIN, ITS PRODUCTION AND USE
;  FILE REFERENCE:  2417US0P
;  CURRENT APPLICATION NUMBER: US/09/242,980
;  CURRENT FILING DATE:  1999-02-28
;  PRIOR APPLICATION NUMBER:  PCT/JP97/03194
;  PRIOR FILING DATE:  1997-08-10
;  PRIOR APPLICATION NUMBER:  JP 8-240880
;  PRIOR FILING DATE:  1996-09-11
;  PRIOR APPLICATION NUMBER:  JP 8-318049
;  PRIOR FILING DATE:  1996-11-28
;  PRIOR APPLICATION NUMBER:  JP 9-135633
;  PRIOR FILING DATE:  1997-05-27
;  NUMBER OF SEQ ID NOS:  35
;  SOFTWARE:
; SEQ ID NO 32
;   LENGTH: 72
;   TYPE: DNA
;   ORGANISM: Human
US-09-242-980-32

  Query Match            59.0%;  Score 11.8;  DB 10;  Length 72;
  Best Local Similarity  86.7%;  Pred. No. 2.7e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GGCGCCCTGTGTCCC 18
           ||||||||| || ||
Db      34 GGCGCCCTGGGTGCC 48


RESULT 9
US-09-873-075A-5/c
; Sequence 5, Application US/09873075A
; Patent No. US20020123123A1
; GENERAL INFORMATION:
;  APPLICANT: Svendsen, Allan
;  APPLICANT:  Schroder Glad, Sanne
;  APPLICANT:  Fukuyama, Shiro
;  APPLICANT:  Matsui, Tomoko
;  TITLE OF INVENTION: Cutinase variants
;  FILE REFERENCE: 10038.200-US
;  CURRENT APPLICATION NUMBER: US/09/873,075A
;  CURRENT FILING DATE:  2001-06-01
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 87
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer
;   NAME/KEY: misc_feature
;   LOCATION: (43)..(43)
;   OTHER INFORMATION: n denotes a, g, c or t
;   NAME/KEY: misc_feature
;   LOCATION: (46)..(46)
;   OTHER INFORMATION: n denotes a, g, c or t
;   NAME/KEY: misc_feature
;   LOCATION: (49)..(49)
;   OTHER INFORMATION: n denotes a, g, c or t
US-09-873-075A-5

  Query Match            58.0%;  Score 11.6;  DB 10;  Length 87;
  Best Local Similarity  70.0%;  Pred. No. 3.4e+03;
  Matches   14;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 AATGGCGCCCTGTGTCCCGA 20
           | :| |||||||  ||| ||
Db      27 ASYGACGCCCTGGATCCAGA 8


RESULT 10
US-09-780-172-54/c
; Sequence 54, Application US/09780172
; Patent No. US20020147163A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Robert McKay
;  APPLICANT:  Susan M. Freier
;  APPLICANT:  Jacqueline Wyatt
;  TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-ALPHA EXPRESSION
;  FILE REFERENCE: RTS-0159
;  CURRENT APPLICATION NUMBER: US/09/780,172
;  CURRENT FILING DATE:  2001-02-08
;  NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 54
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense Oligonucleotide
US-09-780-172-54

  Query Match             57.0%;  Score 11.4;  DB 10;  Length 20;
  Best Local Similarity   92.3%;  Pred. No. 3.8e+03;
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 CCCTGTGTCCCGA 20
           ||||||||| |||
Db      13 CCCTGTGTCACGA 1


RESULT 11
US-09-987-967-3/c
; Sequence 3, Application US/09987967
; Patent No. US20020055144A1
; GENERAL INFORMATION:
;  APPLICANT: Wei et al.
;  TITLE OF INVENTION: Human Hematopoietic - Specific Protein
;  FILE REFERENCE: PF268D1C1
;  CURRENT APPLICATION NUMBER: US/09/987,967
;  CURRENT FILING DATE:  2001-11-16
;  PRIOR APPLICATION NUMBER: PCT/US96/04930
;  PRIOR FILING DATE: 1996-04-11
;  PRIOR APPLICATION NUMBER: 08/837,029
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 09/265,977
;  PRIOR FILING DATE: 1999-03-11
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 28
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: Primer_Bind
;   OTHER INFORMATION: Synthetic primer containing a Bam HI restriction site
;   OTHER INFORMATION: encoding a start AUG, followed by 19 nucleotides of the hHSP
;   OTHER INFORMATION: coding sequence beginning with the first base of the 23rd codon.
US-09-987-967-3

  Query Match             57.0%;  Score 11.4;  DB 10;  Length 28;
  Best Local Similarity   92.3%;  Pred. No. 3.9e+03;
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 AATGGCGCCCTGT 13
           | ||||||||||||
Db      23 AGTGGCGCCCTGT 11


RESULT 12
US-10-046-935-1880/c
; Sequence 1880, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.527C1
;  CURRENT APPLICATION NUMBER: US/10/046,935
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2239
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1880
;   LENGTH: 52
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 18, 19, 20
;   OTHER INFORMATION: n = A,T,C or G
US-10-046-935-1880

  Query Match             57.0%;  Score 11.4;  DB 9;  Length 52;
  Best Local Similarity   92.3%;  Pred. No. 4.1e+03;
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 CGCCCTGTGTCCC 18
           |||||||| ||||
Db      52 CGCCCTGTTTCCC 40


RESULT 13
US-09-864-761-32231
; Sequence 32231, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
```

```
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32231
;   LENGTH: 87
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AF186997.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
;   OTHER INFORMATION: EST_HUMAN HIT: AV761450.1, EVALUE 2.70e-01
US-09-864-761-32231

  Query Match            57.0%;  Score 11.4;  DB 10;  Length 87;
  Best Local Similarity  92.3%;  Pred. No. 4.3e+03;
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 GCCCTGTGTCCCG 19
           |||||| ||||||
Db      25 GCCCTGGGTCCCG 37


RESULT 14
US-09-420-433-34
; Sequence 34, Application US/09420433
; Patent No. US20020098480A1
;  GENERAL INFORMATION:
;    APPLICANT:  Sidransky, David
;    TITLE OF INVENTION:  NUCLEIC ACID MUTATION DETECTION IN
;    TITLE OF INVENTION:  HISTOLOGIC TISSUE
;    NUMBER OF SEQUENCES:  82
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spensley Horn Jubas & Lubitz
;      STREET:  1880 Century Park East, Suite 500
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90067
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/420,433
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/181,664
;      FILING DATE:  JANUARY 14, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wetherell, Jr., Ph.D., John R.
;      REGISTRATION NUMBER:  31,678
;      REFERENCE/DOCKET NUMBER:  PD-3055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 455-5100
;      TELEFAX:  (619) 455-5110
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..16
US-09-420-433-34

  Query Match            56.0%;  Score 11.2;  DB 10;  Length 16;
  Best Local Similarity  81.2%;  Pred. No. 4.7e+03;
  Matches   13;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 GGCGCCCTGTGTCCCG 19
           || |||||||| |||
Db       1 GGTGCCCTGTGAGCCG 16


RESULT 15
US-10-007-805-554
; Sequence 554, Application US/10007805
; Patent No. US20020150581A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Durham, Margarita
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.470C10
;  CURRENT APPLICATION NUMBER: US/10/007,805
;  CURRENT FILING DATE:  2001-12-07
;  NUMBER OF SEQ ID NOS: 593
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 554
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR primer
US-10-007-805-554

  Query Match            56.0%;  Score 11.2;  DB 12;  Length 25;
  Best Local Similarity  81.2%;  Pred. No. 4.9e+03;
  Matches   13;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 GGCGCCCTGTGTCCCG 19
           ||| || || ||||||
Db       4 GGCTCCATGAGTCCCG 19


Search completed: November 23, 2002, 06:42:06
Job time : 18.25 secs
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-7
Perfect score: 20
Sequence: 1 aatggcgccctgtgtcccga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	88	10	AW826045	AW826045 us20f01.x
c	2	13.8	69.0	65	17	AZ922880	AZ922880 SLCot6C05
c	3	13.8	69.0	87	9	AA612425	AA612425 vo05g12.r
	4	13.8	69.0	91	14	W13623	W13623 ma93f03.r1
	5	13.8	69.0	92	9	AA257852	AA257852 MBL2SJ4C8
	6	13.4	67.0	26	17	TA376C02P	AL497551 T. brucei
c	7	13.4	67.0	82	9	AI942604	AI942604 fc73e04.y
	8	13.2	66.0	63	9	AI768372	AI768372 wg84f09.x
c	9	13.2	66.0	73	13	BI256496	BI256496 602977848
c	10	13.2	66.0	97	13	BI764407	BI764407 603046153
	11	13	65.0	65	17	AZ830008	AZ830008 2M0109B05
c	12	13	65.0	85	17	BH865695	BH865695 SALK_0996
	13	13	65.0	92	9	AI465146	AI465146 mv82h01.x
c	14	12.8	64.0	92	17	CNS01ZCG	AL174121 Tetraodon
	15	12.8	64.0	99	17	AZ342558	AZ342558 1M0075L09
	16	12.6	63.0	35	12	BE782279	BE782279 601470033
	17	12.6	63.0	37	12	BF526649	BF526649 602070776
	18	12.6	63.0	37	13	BI257643	BI257643 602967565
	19	12.6	63.0	38	12	BF527175	BF527175 602039857
	20	12.6	63.0	38	12	BF337026	BF337026 602034416
	21	12.6	63.0	41	12	BG164590	BG164590 602342153
c	22	12.6	63.0	46	9	AI370223	AI370223 qu45a02.x
	23	12.6	63.0	48	12	BF341784	BF341784 602016291
	24	12.6	63.0	48	12	BF342690	BF342690 602013643
	25	12.6	63.0	50	12	BF526508	BF526508 602070814
	26	12.6	63.0	50	12	BG261065	BG261065 602372686
	27	12.6	63.0	51	13	BI857448	BI857448 603384413
	28	12.6	63.0	57	12	BG254978	BG254978 602369443
	29	12.6	63.0	57	12	BG391783	BG391783 602418081
	30	12.6	63.0	59	12	BG289259	BG289259 602384267
	31	12.6	63.0	59	13	BI259621	BI259621 602968509
	32	12.6	63.0	61	12	BG258331	BG258331 602379777
	33	12.6	63.0	63	12	BG251491	BG251491 602363952
	34	12.6	63.0	63	13	BI258857	BI258857 602969674
	35	12.6	63.0	65	13	BI259702	BI259702 602971325
	36	12.6	63.0	67	12	BG283508	BG283508 602407805
	37	12.6	63.0	78	12	BG169672	BG169672 602324648
	38	12.6	63.0	79	12	BG112276	BG112276 602282420
	39	12.6	63.0	86	13	BI255929	BI255929 602976561
	40	12.6	63.0	90	12	BF343214	BF343214 602015864
	41	12.6	63.0	97	9	AU076962	AU076962 AU076962
	42	12.4	62.0	90	17	AZ309913	AZ309913 1M0017P17
	43	12.2	61.0	39	14	T49497	T49497 ya76a07.r1
c	44	12.2	61.0	46	9	AI355632	AI355632 qt76h04.x
	45	12.2	61.0	55	14	H23572	H23572 yn67d08.r1

ALIGNMENTS

RESULT 1
AW826045/c
LOCUS AW826045 88 bp mRNA linear EST 17-MAY-2000
DEFINITION us20f01.x1 Soares_NMGBC_B-cell Mus musculus cDNA clone IMAGE:3167641 3' similar to SW:FAN_MOUSE O35242 PROTEIN FAN ;, mRNA sequence.
ACCESSION AW826045
VERSION AW826045.1 GI:7919122
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 88)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: us20f01.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

```
            MGI:1063101
            Trace considered overall poor quality
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .88
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3167641"
                     /clone_lib="Soares_NMGBC_B-cell"
                     /tissue_type="germinal B-cell from resting spleen"
                     /lab_host="DH10B (phage resistant)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCAGGATTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
BASE COUNT       25 a     15 c     22 g     26 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 88;
  Best Local Similarity   84.2%;  Pred. No. 9.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATGGCGCCCTGTGTCCCGA 20
           ||| || ||||||| ||||
Db      84 ATGACGACCTGTGTTCCGA 66


RESULT 2
AZ922880/c
LOCUS       AZ922880                 65 bp    DNA     linear   GSS 07-JUN-2002
DEFINITION  SLCot6C05 Sorghum bicolor SLCot Sorghum bicolor genomic, DNA
            sequence.
ACCESSION   AZ922880
VERSION     AZ922880.1  GI:13433101
KEYWORDS    GSS.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Peterson,D.G., Schulze,S.R., Sciara,E.B., Lee,S.A., Bowers,J.E.,
            Nagel,A., Jiang,N., Tibbitts,D.C., Wessler,S.R. and Paterson,A.H.
  TITLE     Integration of Cot analysis, DNA cloning, and high-throughput
            sequencing facilitates genome characterization and gene discovery
  JOURNAL   Genome Res. 12 (5), 795-807 (2002)
  MEDLINE   21992826
COMMENT     Contact: Peterson DG
            Plant Genome Mapping Laboratory
            University of Georgia
            Room 162, Riverbend Research Bldg., 110 Riverbend Rd., Athens, GA
            30602, USA
            Tel: 706-583-0167
            Fax: 706-583-0160
            Email: dgp@arches.uga.edu
            Class: Hydroxyapatite-fractionated DNA.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Sorghum bicolor"
                     /cultivar="BTx623"
                     /db_xref="taxon:4558"
                     /clone_lib="Sorghum bicolor SLCot"
                     /tissue_type="leaves"
                     /dev_stage="seedling"
                     /note="Vector: pGEM-TA-Easy; A Cot analysis was performed
                     for the sorghum genome. Based on the resulting Cot curve,
                     hydroxyapatite chromatography was used to isolate
                     'highly-repetitive' (HR), 'moderately-repetitive' (MR),
                     and 'single/low-copy' (SL) sequence components from
                     sheared genomic DNA. The three repetition-based DNA
                     components were cloned into E. coli to produce HRCot,
                     MRCot, and SLCot genomic libraries. Blotting and
                     sequencing data indicates that each library is
                     representative of the component from which it was derived.
                     Putative ID listings given for sequences are based on
                     comparison (blastn) with sequences in the NCBI Nr
                     Database. Only the primary match is given (all primary E
                     values are < or = 1.00E-5). In no instance does a 'Cot
                     clone' contain the complete sequence of its putative Nr
                     match."
BASE COUNT       16 a     22 c     13 g     14 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 65;
  Best Local Similarity   88.2%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 ATGGCGCCCTGTGTCCC 18
           ||||||| | |||||||
Db      28 ATGGCGCGCGGTGTCCC 12


RESULT 3
AA612425/c
LOCUS       AA612425                 87 bp    mRNA    linear   EST 01-OCT-1997
DEFINITION  vo05g12.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:1040614 5', mRNA sequence.
ACCESSION   AA612425
VERSION     AA612425.1  GI:2461103
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:580638
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1040614"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       21 a     24 c     29 g     13 t
```

```
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 87;
  Best Local Similarity   88.2%;  Pred. No. 1.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GGCGCCCTGTGTCCCGA 20
           || |||||||||||| |
Db      51 GGAGCCCTGTGTCCCAA 35


RESULT 4
W13623
LOCUS       W13623                   91 bp    mRNA    linear   EST 02-OCT-1997
DEFINITION  ma93f03.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
            IMAGE:318269 5' similar to SW:KRB2_GOATA P02447 KERATIN,
            HIGH-SULFUR MATRIX PROTEIN, IIIB2 ;, mRNA sequence.
ACCESSION   W13623
VERSION     W13623.1  GI:1287661
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI: 208885
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: ETPrimer
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:318269"
                     /clone_lib="Soares mouse p3NMF19.5"
                     /dev_stage="19.5 dpc total fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA was kindly provided by
                     Dr. Minoru Ko (Wayne State University)."
BASE COUNT       19 a     36 c     20 g     16 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 14;  Length 91;
  Best Local Similarity   88.2%;  Pred. No. 1.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GGCGCCCTGTGTCCCGA 20
           || |||||||||||| |
Db      70 GGAGCCCTGTGTCCCAA 86


RESULT 5
AA257852
LOCUS       AA257852                 92 bp    mRNA    linear   EST 22-SEP-1997
DEFINITION  MBL2SJ4C8T3 Brugia malayi second stage larvae JHU96SL-BmL2 Brugia
            malayi cDNA clone L2SJ4C8 5', mRNA sequence.
ACCESSION   AA257852
VERSION     AA257852.1  GI:1893368
KEYWORDS    EST.
SOURCE      Brugia malayi.
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Blaxter,M.L., Waterfall,M., Daub,J., Lizotte,M., Baron,L. and Jones
            ,S.J.
  TITLE     Genes expressed in adult female Brugia malayi
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Blaxter ML
            Institute of Cell, Animal and Population Biology
            University of Edinburgh
            Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
            3JT, UK.
            Tel: +44 131 650 6760
            Fax: +44 131 670 5450
            Email: mark.blaxter@ed.ac.uk
            The ABI trace of this sequence can be viewed at
            http://www.sanger.ac.uk/brugia/L2S/MBL2SJ4C8T3.html
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="Brugia malayi"
                     /strain="TRS Labs"
                     /db_xref="taxon:6279"
                     /clone="L2SJ4C8"
                     /clone_lib="Brugia malayi second stage larvae
                     JHU96SL-BmL2"
                     /sex="mixed"
                     /dev_stage="mosquito derived, second stage larvae (L2)"
                     /lab_host="E. coli XL1-Blue MRF-"
                     /note="Vector: lambdaZapII; Site_1: Eco R I (5' end);
                     Site_2: Xho I (3' end); Brugia malayi is a lymphatic
                     filarial nematode parasite of humans. Full length cDNA was
                     prepared by long-range RT-PCR from mRNA from L2 larvae of
                     the human filarial nematode parasite Brugia malayi using
                     nematode spliced leader (SL, 5'end) and oligo-d(T) (3' end
                     ) primers. The library had an unamplified titre of ~1 x
                     10E6 per ml and ~95% of clones have inserts (mean length
                     ~900 bp). The library is available from The Filarial
                     Genome Project Resource Center: contact Dr. S.A. Williams,
                     Clark Science Center, Smith College, Northampton, MA 01063
                     USA phone +1 413 585 3826 fax  +1 413 585 3786 email
                     genome@smith.edu"
BASE COUNT       20 a     33 c     20 g     19 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 92;
  Best Local Similarity   88.2%;  Pred. No. 1.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GGCGCCCTGTGTCCCGA 20
           | || ||||||||||||
Db       3 GACGGCCTGTGTCCCGA 19


RESULT 6
TA376C02P
LOCUS       TA376C02P                26 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 376c02, forward sequence,
            genomic survey sequence.
```

```
ACCESSION   AL497551
VERSION     AL497551.1  GI:11873273
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nhl@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="376c02"
BASE COUNT        5 a      6 c      6 g      9 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 17;  Length 26;
  Best Local Similarity   93.3%;  Pred. No. 1.4e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 AATGGCGCCCTGTGT 15
           ||||||| ||||||
Db       6 AATGGCGCGCTGTGT 20


RESULT 7
AI942604/c
LOCUS       AI942604                 82 bp    mRNA    linear   EST 07-JUN-2001
DEFINITION  fc73e04.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone
            IMAGE:3727038 5' similar to TR:Q13518 Q13518 PROTEIN KINASE
            C-BINDING PROTEIN RACK8 ;, mRNA sequence.
ACCESSION   AI942604
VERSION     AI942604.1  GI:5707260
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:3727038"
                     /clone_lib="Zebrafish WashU MPIMG EST"
                     /sex="mixed"
                     /tissue_type="26 somite embryos, adult livers, shield
                     stage embryos"
                     /lab_host="XL1-blue MRF"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
                     strand cDNA was primed with a Not I - oligo(dT)15 primer
                     [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT3'];
                     double-stranded cDNA was ligated to Sal I adaptors (BRL),
                     digested with Not I and cloned into the Not I and Sal I
                     sites of the pSPORT1 vector (BRL). Library was constructed
                     by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
                     Institut fuer Molekulare Genetik,Berlin). cDNAs for EST
                     analysis were selected following oligonucleotide
                     hybridization fingerprinting of arrayed clones from
                     zebrafish late somitogenesis (26 ss), adult liver or
                     embryonic shield stage (5.6 h) libraries.  Fingerprint
                     data were used to computationally cluster cDNAs, and a
                     single cDNA from each cluster was chosen for sequencing.
                     In some cases multiple members of the same cluster were
                     sequenced to assess clustering parameters or single clones
                     were sequenced additional times to assess quality
                     control."
BASE COUNT       16 a     24 c     24 g     18 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 82;
  Best Local Similarity   93.3%;  Pred. No. 2.2e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 GCGCCCTGTGTCCCG 19
           ||||||||||||| |
Db      31 GCGCCCTGTGTCCTG 17


RESULT 8
AI768372
LOCUS       AI768372                 63 bp    mRNA    linear   EST 20-DEC-1999
DEFINITION  wg84f09.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:2371817 3' similar to SW:MY18_HUMAN O75293 NEGATIVE
            GROWTH-REGULATORY PROTEIN MYD118 ;, mRNA sequence.
ACCESSION   AI768372
VERSION     AI768372.1  GI:5234881
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 416    Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .63
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2371817"
                      /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                      /lab_host="DH10B"
                      /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                      a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                      Equal amounts of plasmid DNA from five normalized
                      libraries were mixed, and ss circles were made in vitro.
                      Following HAP purification, this DNA was used as tracer in
                      a subtractive hybridization reaction. The driver was
                      PCR-amplified cDNAs from pools of 5,000 clones made from
                      the same 5 libraries. The pools consisted of the following
                      libraries and cloneIDs: Soares NbHSF pool 1:
                      309384-310919, 323208-325895 Soares Nb2HP pool 1:
                      145032-147335, 147720-148103, 148872-149255, 15002 -
                      150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                      758280-760583, 772104-774407 Soares NbHPA pool 1:
                      304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                      pool 1:  723720-726407, 739080-740999 Subtraction by Bento
                      Soares and M. Fatima Bonaldo."
BASE COUNT       14 a     18 c     14 g     17 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 63;
  Best Local Similarity   83.3%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGGCGCCCTGTGTCCCGA 20
           ||| |||||| ||| |||
Db      40 TGGTGCCCTGAGTCTCGA 57


RESULT 9
BI256496/c
LOCUS       BI256496                  73 bp    mRNA    linear   EST 17-JUL-2001
DEFINITION  602977848F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5122810 5',
            mRNA sequence.
ACCESSION   BI256496
VERSION     BI256496.1  GI:14810962
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11299  row: p  column: 11
            High quality sequence stop: 69.
FEATURES              Location/Qualifiers
     source           1. .73
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5122810"
                      /clone_lib="NIH_MGC_12"
                      /tissue_type="cervical carcinoma cell line"
                      /lab_host="DH10B"
                      /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 1.4 kb. Library prepared by Life
                      Technologies."
BASE COUNT       15 a     22 c     24 g     12 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 13;  Length 73;
  Best Local Similarity   83.3%;  Pred. No. 2.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATGGCGCCCTGTGTCCCG 19
           || || ||||||| ||||
Db      57 ATTGCTCCCTGTGACCCG 40


RESULT 10
BI764407/c
LOCUS       BI764407                  97 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603046153F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186413 5',
            mRNA sequence.
ACCESSION   BI764407
VERSION     BI764407.1  GI:15755985
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11465  row: j  column: 14
            High quality sequence stop: 3.
FEATURES              Location/Qualifiers
     source           1. .97
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5186413"
                      /clone_lib="NIH_MGC_116"
                      /lab_host="DH10B"
                      /note="Organ: pooled colon, kidney, stomach; Vector:
                      pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                      source anonymous pool of 3 colons, age 26 yo male, 49 yo
                      female, 71 yo male colon; 46 yo male kidney, and pool of 2
                      stomachs, 62 yo male and 70 yo female.  Library is
                      oligo-dT primed and  directionally cloned (EcoRV site is
                      destroyed upon cloning).  Average insert size 1.4 kb,
                      insert size range 1-3 kb.  Library is normalized and
                      enriched for full-length clones and was constructed by C.
                      Gruber  (Invitrogen).  Research Genetics tracking code
                      023. Note: this is a NIH_MGC Library."
BASE COUNT       21 a     33 c     32 g     11 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 13;  Length 97;
  Best Local Similarity   83.3%;  Pred. No. 2.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 AATGGCGCCCTGTGTCCC 18
```

```
          | |||  ||||||||||||
Db      80 ACTGGGCCCCTGTGTCCC 63


RESULT 11
AZ830008
LOCUS       AZ830008                  65 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0109B05F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0109B05 F, DNA sequence.
ACCESSION   AZ830008
VERSION     AZ830008.1  GI:12999916
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0109  row: B  column: 05
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 65.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0109B05"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       17 a     15 c     15 g     18 t
ORIGIN

  Query Match            65.0%;  Score 13;  DB 17;  Length 65;
  Best Local Similarity  100.0%;  Pred. No. 3.1e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AATGGCGCCCTGT 13
```

```
          ||||||||||||
Db      50 AATGGCGCCCTGT 62


RESULT 12
BH865695/c
LOCUS       BH865695                  85 bp    DNA     linear   GSS 05-AUG-2002
DEFINITION  SALK_099695 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_099695, DNA sequence.
ACCESSION   BH865695
VERSION     BH865695.1  GI:22101593
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated exon of At4g32820.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_099695"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       20 a     19 c     20 g     24 t      2 others
ORIGIN

  Query Match            65.0%;  Score 13;  DB 17;  Length 85;
  Best Local Similarity  100.0%;  Pred. No. 3.4e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AATGGCGCCCTGT 13
          ||||||||||||
Db     31 AATGGCGCCCTGT 19


RESULT 13
AI465146
LOCUS       AI465146                  92 bp    mRNA    linear   EST 09-MAR-1999
DEFINITION  mv82h01.x1 GuayWoodford Beier mouse kidney day 7 Mus musculus cDNA
            clone IMAGE:661585 3', mRNA sequence.
ACCESSION   AI465146
VERSION     AI465146.1  GI:4319176
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
```

```
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end
            Seq primer: Primer name ambiguous.
FEATURES              Location/Qualifiers
     source           1. .92
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:661585"
                      /clone_lib="GuayWoodford Beier mouse kidney day 7"
                      /tissue_type="kidney"
                      /dev_stage="juvenile (7 days old)"
                      /lab_host="SOLR (kanamycin resistant)"
                      /note="Organ: kidney; Vector: pBluescript SK-; Site_1:
                      EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                      Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
                      -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
                      sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3' Library provided
                      by Lisa Guay-Woodford."
BASE COUNT       15 a     22 c     29 g     26 t
ORIGIN

  Query Match             65.0%;  Score 13;  DB 9;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 GCCCTGTGTCCCG 19
           |||||||||||||
Db      41 GCCCTGTGTCCCG 53


RESULT 14
CNS01ZCG/c
LOCUS       CNS01ZCG                  92 bp    DNA     linear   GSS 12-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            220I09 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL174121
VERSION     AL174121.1  GI:7812178
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 92)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 92)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES              Location/Qualifiers
     source           1. .92
                      /organism="Tetraodon nigroviridis"
                      /db_xref="taxon:99883"
                      /clone="220I09"
                      /clone_lib="G"
                      /note="Genoscope sequence ID : C0AG220AE05SP1-end :
                      PUC-Ori"
BASE COUNT       27 a     21 c     42 g      0 t      2 others
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 92;
  Best Local Similarity   87.5%;  Pred. No. 4.4e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TGGCGCCCTGTGTCCC 18
             || |||||| ||||||
Db      33 TGCCGCCCTCTGTCCC 18


RESULT 15
AZ342558
LOCUS       AZ342558                  99 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0075L09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0075L09 R, DNA sequence.
ACCESSION   AZ342558
VERSION     AZ342558.1  GI:10419915
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0075  row: L  column: 09
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 99.
FEATURES              Location/Qualifiers
     source           1. .99
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0075L09"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
```

Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 16 a 31 c 23 g 29 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 17; Length 99;
Best Local Similarity 87.5%; Pred. No. 4.5e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      3 TGGCGCCCTGTGTCCC 18
          ||| |||||||| |||
Db      6 TGGGGCCCTGTGACCC 21
```

Search completed: November 26, 2002, 04:07:25
Job time : 766.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-8
Perfect score: 20
Sequence: 1 gtgcccagccagagcgactg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	15.4	77.0	50	6	AX430850	AX430850 Sequence
2	14.2	71.0	42	6	AX231094	AX231094 Sequence
3	14.2	71.0	42	6	AX231098	AX231098 Sequence
4	14.2	71.0	42	6	AX231555	AX231555 Sequence
5	14.2	71.0	42	6	AX232206	AX232206 Sequence
6	14.2	71.0	42	6	AX233340	AX233340 Sequence
7	14.2	71.0	91	9	HSINOSX01	U65689 Human induc
8	13.4	67.0	51	6	AX157921	AX157921 Sequence
9	13.4	67.0	51	6	AX157922	AX157922 Sequence
10	13.4	67.0	99	11	G34988	G34988 DBX human S
11	13.2	66.0	27	6	AX306740	AX306740 Sequence
c 12	13.2	66.0	27	6	AX306741	AX306741 Sequence
c 13	13.2	66.0	28	6	AR130372	AR130372 Sequence
c 14	13.2	66.0	51	6	AX115349	AX115349 Sequence
c 15	13.2	66.0	76	6	AR121636	AR121636 Sequence
c 16	13.2	66.0	76	6	E55383	E55383 Process for
17	13.2	66.0	90	6	AX306734	AX306734 Sequence
c 18	13.2	66.0	100	11	HSA308722	AJ308722 Homo sapi
c 19	13	65.0	17	6	AR196300	AR196300 Sequence
c 20	12.8	64.0	27	6	I33657	I33657 Sequence 6
c 21	12.8	64.0	30	6	I58430	I58430 Sequence 14
22	12.8	64.0	31	6	I58435	I58435 Sequence 19
c 23	12.8	64.0	55	6	I13710	I13710 Sequence 16
24	12.8	64.0	58	6	AR172735	AR172735 Sequence
25	12.8	64.0	58	6	AR178679	AR178679 Sequence
26	12.8	64.0	58	6	AR181250	AR181250 Sequence
27	12.8	64.0	58	6	AR199931	AR199931 Sequence
c 28	12.8	64.0	60	6	AX097401	AX097401 Sequence
29	12.8	64.0	60	6	E08036	E08036 DNA encodin
30	12.8	64.0	60	6	I24050	I24050 Sequence 14
c 31	12.8	64.0	63	6	A24364	A24364 E.coli OmpA
c 32	12.8	64.0	63	6	A29275	A29275 E.coli ompA
33	12.8	64.0	63	6	A29276	A29276 E.coli ompA
c 34	12.8	64.0	63	6	A42422	A42422 Sequence 4
c 35	12.8	64.0	63	6	A69367	A69367 Sequence 6
c 36	12.8	64.0	63	6	A72044	A72044 Sequence 6
c 37	12.8	64.0	63	6	AR080440	AR080440 Sequence
c 38	12.8	64.0	63	6	AX195143	AX195143 Sequence
c 39	12.8	64.0	63	6	I13659	I13659 Sequence 5
c 40	12.8	64.0	64	6	E08035	E08035 DNA encodin
c 41	12.8	64.0	64	6	I24049	I24049 Sequence 13
c 42	12.8	64.0	65	6	AX329397	AX329397 Sequence
c 43	12.8	64.0	66	6	AX430829	AX430829 Sequence
c 44	12.8	64.0	66	6	AX431269	AX431269 Sequence
c 45	12.8	64.0	66	6	AX431272	AX431272 Sequence

ALIGNMENTS

RESULT 1
AX430850/c
LOCUS AX430850 50 bp mRNA linear PAT 28-JUN-2002
DEFINITION Sequence 4 from Patent WO0240994.
ACCESSION AX430850
VERSION AX430850.1 GI:21655915
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Meshorer,E., Shoham,S., Soreq,H. and Sklan,E.
TITLE System and method for assaying drugs
JOURNAL Patent: WO 0240994-A 4 23-MAY-2002;

```
          YISSUM RES DEV CO (IL)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        5 a     17 c     16 g     12 t
ORIGIN

  Query Match             77.0%;  Score 15.4;  DB 6;  Length 50;
  Best Local Similarity   94.1%;  Pred. No. 6.3e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 TGCCCAGCCAGAGCGAC 18
           | ||||||||||||||||
Db      50 TCCCCAGCCAGAGCGAC 34


RESULT 2
AX231094
LOCUS       AX231094                  42 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 490 from Patent WO0162787.
ACCESSION   AX231094
VERSION     AX231094.1  GI:15592450
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Herath,H.M., Parekh,R.B., Rohlff,C.O., Terrett,J.A. and Tyson,K.L.
  TITLE     Diagnosis and treatment of bipolar affective disorder
  JOURNAL   Patent: WO 0162787-A 490 30-AUG-2001;
            Oxford GlycoSciences (UK) Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             <1. .>42
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC69698.1"
                     /db_xref="GI:15592451"
                     /translation="EWVAIESDSVQPVP"
BASE COUNT        6 a     12 c     15 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 42;
  Best Local Similarity   84.2%;  Pred. No. 2.5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCCAGCCAGAGCGACT 19
           ||| ||| | |||||||||
Db       7 GTGGCCATCGAGAGCGACT 25


RESULT 3
AX231098
LOCUS       AX231098                  42 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 494 from Patent WO0162787.
ACCESSION   AX231098
VERSION     AX231098.1  GI:15592454
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Herath,H.M., Parekh,R.B., Rohlff,C.O., Terrett,J.A. and Tyson,K.L.
  TITLE     Diagnosis and treatment of bipolar affective disorder
  JOURNAL   Patent: WO 0162787-A 494 30-AUG-2001;
            Oxford GlycoSciences (UK) Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        6 a     12 c     15 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 42;
  Best Local Similarity   84.2%;  Pred. No. 2.5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCCAGCCAGAGCGACT 19
           ||| ||| | |||||||||
Db       7 GTGGCCATCGAGAGCGACT 25


RESULT 4
AX231555
LOCUS       AX231555                  42 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 453 from Patent WO0163294.
ACCESSION   AX231555
VERSION     AX231555.1  GI:15592461
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Herath,H.M., Parekh,R.B. and Rohlff,C.
  TITLE     Diagnosis of bipolar affective disorder (bad) and unipolar
            depression
  JOURNAL   Patent: WO 0163294-A 453 30-AUG-2001;
            Oxford GlycoSciences (UK) Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             <1. .>42
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC69700.1"
                     /db_xref="GI:15592462"
                     /translation="EWVAIESDSVQPVP"
BASE COUNT        6 a     12 c     15 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 42;
  Best Local Similarity   84.2%;  Pred. No. 2.5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCCAGCCAGAGCGACT 19
           ||| ||| | |||||||||
Db       7 GTGGCCATCGAGAGCGACT 25


RESULT 5
AX232206
LOCUS       AX232206                  42 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 484 from Patent WO0163293.
ACCESSION   AX232206
VERSION     AX232206.1  GI:15592539
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Herath,M.A., Parekh,R.B. and Rohlff,C.O.
  TITLE     Diagnosis and treatment of schizophrenia
  JOURNAL   Patent: WO 0163293-A 484 30-AUG-2001;
            Oxford GlycoSciences (UK) Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
     CDS             <1. .>42
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC69707.1"
                     /db_xref="GI:15592540"
                     /translation="EWVAIESDSVQPVP"
BASE COUNT        6 a     12 c     15 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 42;
  Best Local Similarity   84.2%;  Pred. No. 2.5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCCCAGCCAGAGCGACT 19
          ||| ||| | |||||||||
Db      7 GTGGCCATCGAGAGCGACT 25


RESULT 6
AX233340
LOCUS       AX233340                  42 bp    DNA     linear   PAT 11-SEP-2001
DEFINITION  Sequence 664 from Patent WO0162785.
ACCESSION   AX233340
VERSION     AX233340.1  GI:15592679
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Herath,H.M., Parekh,R.B., Rohlff,C.C., Terrett,J.A. and Tyson,K.L.
  TITLE     Protein and gene and their use for diagnosis and treatment of
            schizophrenia
  JOURNAL   Patent: WO 0162785-A 664 30-AUG-2001;
            Oxford GlycoSciences (UK) Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             <1. .>42
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC69730.1"
                     /db_xref="GI:15592680"
                     /translation="EWVAIESDSVQPVP"
BASE COUNT        6 a     12 c     15 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 42;
  Best Local Similarity   84.2%;  Pred. No. 2.5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCCCAGCCAGAGCGACT 19
          ||| ||| | |||||||||
Db      7 GTGGCCATCGAGAGCGACT 25


RESULT 7
HSINOSX01
LOCUS       HSINOSX01                 91 bp    DNA     linear   PRI 27-JAN-1998
DEFINITION  Human inducible nitric oxide synthase (iNOS) pseudogene, exon 1.
ACCESSION   U65689
VERSION     U65689.1  GI:1621317
KEYWORDS    .
SEGMENT     1 of 11
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Park,C.S., Park,R., Lee,H.S., Lee,H.Y. and Krishna,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-1996) Section of Chemical Pharmacology, National
            Institutes of Health, Bethesda, MD 20892, USA
REFERENCE   2  (bases 1 to 91)
  AUTHORS   Park,C.S., Lee,H.S., Lee,H.Y. and Krishna,G.
  TITLE     An unprocessed pseudogene of inducible nitric oxide synthase gene
            in human
  JOURNAL   Nitric Oxide 1 (4), 294-300 (1997)
  MEDLINE   98368274
   PUBMED   9441901
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17q11.2"
     exon            <1. .29
                     /gene="iNOS"
                     /number=1
                     /pseudo
     intron          30. .91
                     /gene="iNOS"
BASE COUNT       22 a     29 c     25 g     15 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 91;
  Best Local Similarity   84.2%;  Pred. No. 2.4e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCCCAGCCAGAGCGACT 19
          || ||||||||||||  ||
Db     34 GTTCCCAGCCAGAGCTGCT 52


RESULT 8
AX157921
LOCUS       AX157921                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 1249 from Patent WO0140521.
ACCESSION   AX157921
VERSION     AX157921.1  GI:14539252
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 1249 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (1250 is other entry)
                     Accession number cg28464080"
BASE COUNT       10 a      8 c     20 g     13 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 51;
  Best Local Similarity   93.3%;  Pred. No. 6.1e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 CAGCCAGAGCGACTG 20
          ||||||||| |||||
Db      9 CAGCCAGAGAGACTG 23


RESULT 9
AX157922
LOCUS       AX157922                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 1250 from Patent WO0140521.
ACCESSION   AX157922
```

```
VERSION     AX157922.1  GI:14539253
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 1250 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (1249 is other entry)
                     Accession number cg28464080"
BASE COUNT       10 a      9 c     19 g     13 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 51;
  Best Local Similarity   93.3%;  Pred. No. 6.1e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 CAGCCAGAGCGACTG 20
          ||||||||| |||||
Db      9 CAGCCAGAGAGACTG 23


RESULT 10
G34988
LOCUS       G34988                     99 bp    DNA     linear   STS 06-DEC-1999
DEFINITION  DBX human STS derived from X-linked gene Homo sapiens STS genomic,
            sequence tagged site.
ACCESSION   G34988
VERSION     G34988.1  GI:2599270
KEYWORDS    STS.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Lahn,B.T. and Page,D.C.
  TITLE     Functional coherence of the human Y chromosome
  JOURNAL   Science 278 (5338), 675-680 (1997)
  MEDLINE   98022381
   PUBMED   9381176
COMMENT
            Contact: Bruce T. Lahn
            Page Lab
            Whitehead Institue for Biomedical Research
            9 Cambridge Center, Cambridge, MA 02142, USA
            Tel: 6172585164
            Email: blahn@wi.mit.edu
            Primer A: CTACATGCAGATGACATGGTG
            Primer B: GGCCAAGGTGCATAGGTG
            STS size: 99
            PCR Profile:
                  Denaturation: 94 for 1 min
                  Annealing: 60 for 45 sec
                  Polymerization: 72 for 45 sec
                  PCR Cycles: 30
                  Thermal Cycler: MJ Research
            Protocol:
                  Template:    50 ng
                  Primer:      each 1 uM
                  dNTPs:       100 uM
                  Taq Polymerase:  1 unit
                  Total Vol:   20 ul

            Buffer:
                  MgCl2:    1.5 mM
                  KCl:      50 mM
                  Tris pH 9: 10 mM
                  Triton X-100: 0.1%.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="human STS derived from X-linked gene"
                     /note="human STS derived from the cDNA sequence of
                     X-linked gene"
     STS             1. .99
     primer_bind     1. .21
     primer_bind     complement(82. .99)
BASE COUNT       25 a     29 c     19 g     26 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 11;  Length 99;
  Best Local Similarity   93.3%;  Pred. No. 5.9e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GTGCCCAGCCAGAGC 15
          |||||||||| ||||
Db     28 GTGCCCAGCCTGAGC 42


RESULT 11
AX306740
LOCUS       AX306740                   27 bp    DNA     linear   PAT 11-DEC-2001
DEFINITION  Sequence 58 from Patent WO0187925.
ACCESSION   AX306740
VERSION     AX306740.1  GI:17645907
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Rosendahl,M.S., Cox,G.N. and Doherty,D.H.
  TITLE     Methods for refolding proteins containing free cysteine residues
  JOURNAL   Patent: WO 0187925-A 58 22-NOV-2001;
            Bolder Biotechnology, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer"
BASE COUNT        6 a     11 c      6 g      4 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 27;
  Best Local Similarity   83.3%;  Pred. No. 7.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGCCCAGCCAGAGCGACT 19
          ||| ||||||  ||||||
Db      2 TGCACAGCCACTGCGACT 19


RESULT 12
AX306741/c
LOCUS       AX306741                   27 bp    DNA     linear   PAT 11-DEC-2001
DEFINITION  Sequence 59 from Patent WO0187925.
ACCESSION   AX306741
VERSION     AX306741.1  GI:17645908
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Rosendahl,M.S., Cox,G.N. and Doherty,D.H.
  TITLE     Methods for refolding proteins containing free cysteine residues
  JOURNAL   Patent: WO 0187925-A 59 22-NOV-2001;
            Bolder Biotechnology, Inc. (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer"
BASE COUNT        4 a      6 c     11 g      6 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 27;
  Best Local Similarity   83.3%;  Pred. No. 7.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGCCCAGCCAGAGCGACT 19
          ||| ||||||  ||||||
Db     26 TGCACAGCCACTGCGACT 9


RESULT 13
AR130372/c
LOCUS       AR130372                  28 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 2 from patent US 6190657.
ACCESSION   AR130372
VERSION     AR130372.1  GI:14118697
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Pawelek,J.M., Bermudes,D. and Low,K.Brooks.
  TITLE     Vectors for the diagnosis and treatment of solid tumors including
            melanoma
  JOURNAL   Patent: US 6190657-A 2 20-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unknown"
BASE COUNT        6 a      6 c     10 g      6 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 28;
  Best Local Similarity   83.3%;  Pred. No. 7.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 GCCCAGCCAGAGCGACTG 20
          |||| |||| ||| ||||
Db     28 GCCCTGCCATAGCCACTG 11


RESULT 14
AX115349/c
LOCUS       AX115349                  51 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 472 from Patent WO0129262.
ACCESSION   AX115349
VERSION     AX115349.1  GI:14032291
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Picoult-Newburg,L. and Pohl,M.
  TITLE     Genotyping reagents, kits and methods of use thereof
  JOURNAL   Patent: WO 0129262-A 472 26-APR-2001;
            Orchid BioSciences, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       16 a      7 c     16 g     12 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 7.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGCCCAGCCAGAGCGACT 19
          |||| ||| |||| ||||
Db     48 TGCCTAGCAAGAGTGACT 31


RESULT 15
AR121636/c
LOCUS       AR121636                  76 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 31 from patent US 6160089.
ACCESSION   AR121636
VERSION     AR121636.1  GI:14105212
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Honjo,M., Naitoh,N., Uchida,H., Mochizuki,D. and Matsumoto,K.
  TITLE     Method for secretory production of human growth hormone
  JOURNAL   Patent: US 6160089-A 31 12-DEC-2000;
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="unknown"
BASE COUNT       13 a     23 c     22 g     18 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 76;
  Best Local Similarity   83.3%;  Pred. No. 7.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 GCCCAGCCAGAGCGACTG 20
          || |||||||| || |||
Db     25 GCGCAGCCAGAACGCCTG 8


Search completed: December  3, 2002, 18:12:49
Job time : 309.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-8
Perfect score: 20
Sequence: 1 gtgcccagccagagcgactg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	13.2	66.0	28	4	US-08-658-034-2	Sequence 2, Appli
c	2	13.2	66.0	76	4	US-09-348-578-31	Sequence 31, Appl
c	3	13.2	66.0	76	4	US-09-699-684-31	Sequence 31, Appl
c	4	13	65.0	17	4	US-08-679-645-765	Sequence 765, App
c	5	12.8	64.0	27	1	US-07-749-446-6	Sequence 6, Appli
c	6	12.8	64.0	30	1	US-08-360-841-14	Sequence 14, Appl
	7	12.8	64.0	31	1	US-08-360-841-19	Sequence 19, Appl
c	8	12.8	64.0	55	1	US-07-826-928A-16	Sequence 16, Appl
	9	12.8	64.0	58	4	US-09-339-913B-54	Sequence 54, Appl
	10	12.8	64.0	58	4	US-09-339-904A-54	Sequence 54, Appl
	11	12.8	64.0	58	4	US-08-769-062B-54	Sequence 54, Appl
	12	12.8	64.0	58	4	US-09-344-002B-54	Sequence 54, Appl
	13	12.8	64.0	58	4	US-09-559-565C-54	Sequence 54, Appl
	14	12.8	64.0	60	1	US-08-325-243A-14	Sequence 14, Appl
c	15	12.8	64.0	63	1	US-07-842-089E-5	Sequence 5, Appli
c	16	12.8	64.0	63	1	US-08-264-485-5	Sequence 5, Appli
c	17	12.8	64.0	63	2	US-09-003-081-7	Sequence 7, Appli
c	18	12.8	64.0	63	3	US-08-648-262-7	Sequence 7, Appli
c	19	12.8	64.0	63	3	US-08-648-263-7	Sequence 7, Appli
c	20	12.8	64.0	64	1	US-08-325-243A-13	Sequence 13, Appl
c	21	12.8	64.0	69	1	US-08-215-138-18	Sequence 18, Appl
c	22	12.8	64.0	69	1	US-08-215-138-19	Sequence 19, Appl
c	23	12.8	64.0	69	1	US-08-407-344-18	Sequence 18, Appl
c	24	12.8	64.0	69	1	US-08-407-344-19	Sequence 19, Appl
c	25	12.8	64.0	72	1	US-07-826-928A-19	Sequence 19, Appl
c	26	12.8	64.0	73	1	US-07-854-845B-2	Sequence 2, Appli
c	27	12.8	64.0	77	3	US-08-666-354A-10	Sequence 10, Appl
c	28	12.8	64.0	78	4	US-08-976-183A-1	Sequence 1, Appli
c	29	12.8	64.0	78	4	US-08-976-183A-3	Sequence 3, Appli
c	30	12.8	64.0	87	1	US-07-826-928A-26	Sequence 26, Appl
	31	12.8	64.0	87	1	US-07-826-928A-27	Sequence 27, Appl
c	32	12.8	64.0	87	2	US-08-449-287-18	Sequence 18, Appl
	33	12.8	64.0	87	2	US-08-449-287-19	Sequence 19, Appl
c	34	12.8	64.0	93	1	US-08-458-120-3	Sequence 3, Appli
c	35	12.8	64.0	93	2	US-08-187-186A-3	Sequence 3, Appli
c	36	12.8	64.0	93	2	US-08-442-497C-3	Sequence 3, Appli
c	37	12.8	64.0	93	4	US-09-333-033-3	Sequence 3, Appli
c	38	12.6	63.0	20	3	US-08-872-855-12	Sequence 12, Appl
	39	12.6	63.0	37	3	US-08-862-431-47	Sequence 47, Appl
c	40	12.6	63.0	63	1	US-08-590-804-31	Sequence 31, Appl
c	41	12.4	62.0	17	4	US-09-610-155D-3	Sequence 3, Appli
	42	12.4	62.0	30	4	US-09-142-334-2	Sequence 2, Appli
c	43	12.4	62.0	30	4	US-09-142-334-5	Sequence 5, Appli
c	44	12.4	62.0	60	1	US-08-484-192-154	Sequence 154, App
c	45	12.4	62.0	60	1	US-08-484-192-156	Sequence 156, App

ALIGNMENTS

RESULT 1
US-08-658-034-2/c
; Sequence 2, Application US/08658034
; Patent No. 6190657
; GENERAL INFORMATION:
; APPLICANT: Pawelek, John M.
; APPLICANT: Bermudes, David
; APPLICANT: Low, Kenneth Brooks
; TITLE OF INVENTION: VECTORS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF SOLID TUMORS INCLUDING MELANOMA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,034
; FILING DATE: On Even Date Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Baldwin, Geraldine F.
; REGISTRATION NUMBER: 31,232
; REFERENCE/DOCKET NUMBER: 8002-036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-658-034-2

Query Match 66.0%; Score 13.2; DB 4; Length 28;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GCCCAGCCAGAGCGACTG 20
|||| |||| ||| ||||

```
Db      28 GCCCTGCCATAGCCACTG 11


RESULT 2
US-09-348-578-31/c
; Sequence 31, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
;  APPLICANT: HONJO, Masaru
;  APPLICANT: NAITOH, Naokazu
;  APPLICANT: UCHIDA, Hiroshi
;  APPLICANT: MOCHIZUKI, Daisuke
;  APPLICANT: MATSUMOTO, Kazuya
;  TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
;  FILE REFERENCE: 029430-421
;  CURRENT APPLICATION NUMBER: US/09/348,578
;  CURRENT FILING DATE: 1999-07-07
;  EARLIER APPLICATION NUMBER: JP 193003/1998
;  EARLIER FILING DATE: 1998-07-08
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
;  LENGTH: 76
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  LOCATION: (1)..(76)
;  OTHER INFORMATION: Description of Artificial Sequence:Designed DNA sequence to
;  OTHER INFORMATION: act as a primer for PCR to produce the latter half of OppA
;  OTHER INFORMATION: secretion signal
US-09-348-578-31

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 76;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCCCAGCCAGAGCGACTG 20
           || |||||||| || |||
Db      25 GCGCAGCCAGAACGCCTG 8


RESULT 3
US-09-699-684-31/c
; Sequence 31, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
;  APPLICANT: HONJO, Masaru
;  APPLICANT:  NAITOH, Naokazu
;  APPLICANT:  UCHIDA, Hiroshi
;  APPLICANT:  MOCHIZUKI, Daisuke
;  APPLICANT:  MATSUMOTO, Kazuya
;  TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
;  FILE REFERENCE: 029430-421
;  CURRENT APPLICATION NUMBER: US/09/699,684
;  CURRENT FILING DATE:  2000-10-31
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
;  LENGTH: 76
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  LOCATION: (1)..(76)
;  OTHER INFORMATION: Description of Artificial Sequence:Designed DNA sequence to
;  OTHER INFORMATION: act as a primer for PCR to produce the latter half of OppA
;  OTHER INFORMATION: secretion signal
US-09-699-684-31

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 76;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCCCAGCCAGAGCGACTG 20
           || |||||||| || |||
Db      25 GCGCAGCCAGAACGCCTG 8


RESULT 4
US-08-679-645-765/c
; Sequence 765, Application US/08679645
; Patent No. 6350934
;  GENERAL INFORMATION:
;    APPLICANT:  Zwick, Michael G.
;    APPLICANT:  Edington, Brent E.
;    APPLICANT:  McSwiggen, James A.
;    APPLICANT:  Merlo, Patricia Ann Owens
;    APPLICANT:  Guo, Lining
;    APPLICANT:  Skokut, Thomas A.
;    APPLICANT:  Young, Scott A.
;    APPLICANT:  Folkerts, Otto
;    APPLICANT:  Merlo, Donald J.
;    TITLE OF INVENTION:  COMPOSITION AND METHODS FOR
;    TITLE OF INVENTION:  MODULATION OF GENE EXPRESSION
;    TITLE OF INVENTION:  IN PLANTS
;    NUMBER OF SEQUENCES:  1263
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/679,645
;      FILING DATE:  July 12, 1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/001,135
;      FILING DATE:  July 13, 1995
;      APPLICATION NUMBER:  08/300,726
;      FILING DATE:  September 2, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  219/247
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  765:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-679-645-765

  Query Match             65.0%;  Score 13;  DB 4;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+03;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 CAGCCAGAGCGAC 18
           |||||||||||||
Db      15 CAGCCAGAGCGAC 3


RESULT 5
```

```
US-07-749-446-6/c
; Sequence 6, Application US/07749446
; Patent No. 5593857
;  GENERAL INFORMATION:
;    APPLICANT:  Higaki, Jeffrey N.
;    APPLICANT:  Tischer, Edmund G.
;    APPLICANT:  Cordell, Barbara
;    APPLICANT:  Thompson, Stewart A.
;    TITLE OF INVENTION:  PRODUCTION OF HOMOGENEOUS CILIARY
;    TITLE OF INVENTION:  NEUROTROPHIC FACTOR
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  California Biotechnology Inc.
;      STREET:  2450 Bayshore Parkway
;      CITY:  Mountain View
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94043
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/749,446
;      FILING DATE:  19911008
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shearer, Peter R.
;      REGISTRATION NUMBER:  28,117
;      REFERENCE/DOCKET NUMBER:  PC43:US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-962-5860
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-07-749-446-6

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 27;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CCCAGCCAGAGCGACT 19
           || || ||||||||||
Db      27 CCGAGGCAGAGCGACT 12


RESULT 6
US-08-360-841-14/c
; Sequence 14, Application US/08360841
; Patent No. 5652120
;  GENERAL INFORMATION:
;    APPLICANT:  PARK, Seung Kook
;    APPLICANT:  LEE, Kang Moon
;    APPLICANT:  NHO, Kyoo Seung
;    APPLICANT:  KOH, Yeo Wook
;    APPLICANT:  KWON, Chang Hyuk
;    APPLICANT:  CHUNG, Ju Young
;    APPLICANT:  JEE, Young Su
;    APPLICANT:  Yu, Young Hyo
;    TITLE OF INVENTION:  A No. 5652120el Gene Coding Human Epidermal
;    TITLE OF INVENTION:  Growth Factor and Process for Preparing the Same
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Darby & Darby PC
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  US
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/360,841
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, S. Peter
;      REGISTRATION NUMBER:  25,351
;      REFERENCE/DOCKET NUMBER:  0136/0A760
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7700
;      TELEFAX:  212-753-6237
;      TELEX:  236687
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      CLONE:  OMP C2 PRIMER
US-08-360-841-14

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 30;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           |||||||| || ||||
Db      20 CCAGCCAGTGCCACTG 5


RESULT 7
US-08-360-841-19
; Sequence 19, Application US/08360841
; Patent No. 5652120
;  GENERAL INFORMATION:
;    APPLICANT:  PARK, Seung Kook
;    APPLICANT:  LEE, Kang Moon
;    APPLICANT:  NHO, Kyoo Seung
;    APPLICANT:  KOH, Yeo Wook
;    APPLICANT:  KWON, Chang Hyuk
;    APPLICANT:  CHUNG, Ju Young
;    APPLICANT:  JEE, Young Su
;    APPLICANT:  Yu, Young Hyo
;    TITLE OF INVENTION:  A No. 5652120el Gene Coding Human Epidermal
;    TITLE OF INVENTION:  Growth Factor and Process for Preparing the Same
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Darby & Darby PC
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  US
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/360,841
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME:  Ludwig, S. Peter
;       REGISTRATION NUMBER:  25,351
;       REFERENCE/DOCKET NUMBER:  0136/0A760
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-527-7700
;       TELEFAX:  212-753-6237
;       TELEX:  236687
;   INFORMATION FOR SEQ ID NO:  19:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  31 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA to mRNA
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;     IMMEDIATE SOURCE:
;       CLONE:  OMP N3 PRIMER
US-08-360-841-19

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 31;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           |||||||| || ||||
Db      15 CCAGCCAGTGCCACTG 30


RESULT 8
US-07-826-928A-16/c
; Sequence 16, Application US/07826928A
; Patent No. 5439829
;  GENERAL INFORMATION:
;     APPLICANT:  Anderson, Leslie D.
;     APPLICANT:  Cook, James A.
;     APPLICANT:  David, Gary S.
;     APPLICANT:  Hochschwender, Susan M.
;     APPLICANT:  Kasher, Mary S.
;     APPLICANT:  Smith, Michele C.
;     APPLICANT:  Stemmer, Willem P.
;     TITLE OF INVENTION:  METHOD OF IMMOBILIZING AND CROSS LINKING
;     TITLE OF INVENTION:  PROTEINS AND OTHER MOLECULES AND USES THEREOF
;     NUMBER OF SEQUENCES:  38
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Eli Lilly and Company
;       STREET:  Lilly Corporate Center
;       CITY:  Indianapolis
;       STATE:  IN
;       COUNTRY:  USA
;       ZIP:  46285
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/826,928A
;       FILING DATE:  19920124
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Murphy, Richard B.
;       REGISTRATION NUMBER:  35,296
;       REFERENCE/DOCKET NUMBER:  X8180A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  317-276-3589
;       TELEFAX:  317-276-1294
;   INFORMATION FOR SEQ ID NO:  16:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  55 base pairs
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-07-826-928A-16

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 55;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           |||||||| || ||||
Db      26 CCAGCCAGTGCCACTG 11


RESULT 9
US-09-339-913B-54
; Sequence 54, Application US/09339913B
; Patent No. 6303344
; GENERAL INFORMATION:
;  APPLICANT: Patten, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYPEPTIDE ENGINEERING
;  FILE REFERENCE: 02-020503US
;  CURRENT APPLICATION NUMBER: US/09/339,913B
;  CURRENT FILING DATE:  1999-06-24
;  PRIOR APPLICATION NUMBER: 08/769,062
;  PRIOR FILING DATE: 1996-12-18
;  PRIOR APPLICATION NUMBER: 08/198,431
;  PRIOR FILING DATE: 1994-02-17
;  PRIOR APPLICATION NUMBER: 08/425,684
;  PRIOR FILING DATE: 1995-04-18
;  PRIOR APPLICATION NUMBER: 08/537,874
;  PRIOR FILING DATE: 1995-10-30
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: degenerate
;   OTHER INFORMATION: oligonucleotide used for codon usage library
US-09-339-913B-54

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 58;
  Best Local Similarity   81.2%;  Pred. No. 1.8e+03;
  Matches   13;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           ||||||||:|| ||:|
Db      15 CCAGCCAGHGCHACHG 30


RESULT 10
US-09-339-904A-54
; Sequence 54, Application US/09339904A
; Patent No. 6319713
; GENERAL INFORMATION:
;  APPLICANT: Patten, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYPEPTIDE ENGINEERING
;  FILE REFERENCE: 02-020504US
;  CURRENT APPLICATION NUMBER: US/09/339,904A
;  CURRENT FILING DATE:  1999-06-24
;  PRIOR APPLICATION NUMBER: 08/769,062
;  PRIOR FILING DATE: 1996-12-18
;  PRIOR APPLICATION NUMBER: 08/198,431
;  PRIOR FILING DATE: 1994-02-17
;  PRIOR APPLICATION NUMBER: 08/425,684
;  PRIOR FILING DATE: 1995-04-18
;  PRIOR APPLICATION NUMBER: 08/537,874
;  PRIOR FILING DATE: 1995-10-30
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
```

```
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: degenerate
;   OTHER INFORMATION: oligonucleotide used for codon usage library
US-09-339-904A-54

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 58;
  Best Local Similarity   81.2%;  Pred. No. 1.8e+03;
  Matches   13;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           |||||||||:|| ||:|
Db      15 CCAGCCAGHGCHACHG 30


RESULT 11
US-08-769-062B-54
; Sequence 54, Application US/08769062B
; Patent No. 6335160
; GENERAL INFORMATION:
;  APPLICANT: Patten, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYPEPTIDE ENGINEERING
;  FILE REFERENCE: 02-020500US
;  CURRENT APPLICATION NUMBER: US/08/769,062B
;  CURRENT FILING DATE:  1996-12-18
;  PRIOR APPLICATION NUMBER: 08/198,431
;  PRIOR FILING DATE: 1994-02-17
;  PRIOR APPLICATION NUMBER: 08/425,684
;  PRIOR FILING DATE: 1995-04-18
;  PRIOR APPLICATION NUMBER: 08/537,874
;  PRIOR FILING DATE: 1995-10-30
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: degenerate
;   OTHER INFORMATION: oligonucleotide used for codon usage library
US-08-769-062B-54

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 58;
  Best Local Similarity   81.2%;  Pred. No. 1.8e+03;
  Matches   13;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           |||||||||:|| ||:|
Db      15 CCAGCCAGHGCHACHG 30


RESULT 12
US-09-344-002B-54
; Sequence 54, Application US/09344002B
; Patent No. 6355484
; GENERAL INFORMATION:
;  APPLICANT: Patten, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYPEPTIDE ENGINEERING
;  FILE REFERENCE: 02-020502US
;  CURRENT APPLICATION NUMBER: US/09/344,002B
;  CURRENT FILING DATE:  1999-06-24
;  PRIOR APPLICATION NUMBER: 08/769,062
;  PRIOR FILING DATE: 1996-12-18
;  PRIOR APPLICATION NUMBER: 08/198,431
;  PRIOR FILING DATE: 1994-02-17
;  PRIOR APPLICATION NUMBER: 08/425,684
;  PRIOR FILING DATE: 1995-04-18
;  PRIOR APPLICATION NUMBER: 08/537,874
;  PRIOR FILING DATE: 1995-10-30
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: degenerate
;   OTHER INFORMATION: oligonucleotide used for codon usage library
US-09-344-002B-54

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 58;
  Best Local Similarity   81.2%;  Pred. No. 1.8e+03;
  Matches   13;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           |||||||||:|| ||:|
Db      15 CCAGCCAGHGCHACHG 30


RESULT 13
US-09-559-565C-54
; Sequence 54, Application US/09559565C
; Patent No. 6455253
; GENERAL INFORMATION:
;  APPLICANT: Patten, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYPEPTIDE ENGINEERING
;  FILE REFERENCE: 02-020506US
;  CURRENT APPLICATION NUMBER: US/09/559,565C
;  CURRENT FILING DATE:  2000-04-27
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: degenerate
;   OTHER INFORMATION: oligonucleotide used for codon usage library
US-09-559-565C-54

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 58;
  Best Local Similarity   81.2%;  Pred. No. 1.8e+03;
  Matches   13;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           |||||||||:|| ||:|
Db      15 CCAGCCAGHGCHACHG 30


RESULT 14
US-08-325-243A-14
; Sequence 14, Application US/08325243A
; Patent No. 5541288
;  GENERAL INFORMATION:
;    APPLICANT:  Nakano, Shigeru
;    APPLICANT:  Mabuchi, Toshiyuki
;    APPLICANT:  Tada, Miki
;    APPLICANT:  Taoda, Yasuo
;    APPLICANT:  Sugino, Dan
;    APPLICANT:  Kono, Yoshio
;    APPLICANT:  Nishimura, Kaoru
;    APPLICANT:  Okushima, Minoru
;    TITLE OF INVENTION:  No. 5541288el Peptide Having Elastase Inhibitory
;    TITLE OF INVENTION:  Activity and Producing Method Thereof
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illlinois
;      COUNTRY:  United States of America
```

```
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/325,243A
;       FILING DATE:  16-DEC-94
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 5-31758
;       FILING DATE:  22-FEB-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 5541288and, Greta E.
;       REGISTRATION NUMBER:  35,302
;       REFERENCE/DOCKET NUMBER:  19036/32311
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/474-6300
;       TELEFAX:  312/474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  14:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  60 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Other nucleic acid, Synthesized DNA
;     FEATURE:
;       NAME/KEY:  sig peptide
;       LOCATION:  1..60
;       IDENTIFICATION METHOD:  S
;       OTHER INFORMATION:  Complementary Strand is formed with Sequence
;       OTHER INFORMATION:  of SEQ. ID. NO.13.
US-08-325-243A-14

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 60;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           |||||||||||  || |
Db      23 CCAGCCAGAGCAACAG 38


RESULT 15
US-07-842-089E-5/c
; Sequence 5, Application US/07842089E
; Patent No. 5356875
;  GENERAL INFORMATION:
;    APPLICANT:  SARMIENTOS, PAOLO
;    APPLICANT:  DE TAXIS DU POET, PHILIPPE
;    APPLICANT:  NITTI, GIAMPAOLO
;    APPLICANT:  SCACHERI, EMANUELA
;    TITLE OF INVENTION:  ANTI-THROMBIN POLYPEPTIDES
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 Jefferson Davis Highway, Fourth Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/842,089E
;      FILING DATE:  26-FEB-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5356875man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  769-265-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)412-3000
;      TELEFAX:  (703)413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  63 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Escherichia coli
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..63
;      OTHER INFORMATION:  /function= "leader peptide"
;      OTHER INFORMATION:  /standard_name= "OmpA leader"
US-07-842-089E-5

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 63;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           ||||||||  ||  ||||
Db      41 CCAGCCAGTGCCACTG 26


Search completed: November 23, 2002, 06:36:08
Job time : 23.55 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-8
Perfect score: 20
Sequence: 1 gtgcccagccagagcgactg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	14.2	71.0	42	10	US-09-791-378-664	Sequence 664, App
c 2	13.4	67.0	100	10	US-09-864-761-18482	Sequence 18482, A
3	12.8	64.0	58	10	US-09-559-671A-54	Sequence 54, Appl
c 4	12.8	64.0	63	10	US-09-760-008A-4	Sequence 4, Appli
c 5	12.8	64.0	65	10	US-09-875-494-20	Sequence 20, Appl
c 6	12.8	64.0	93	12	US-10-004-832-3	Sequence 3, Appli
c 7	12.8	64.0	99	10	US-09-815-343-1484	Sequence 1484, Ap
c 8	12.6	63.0	79	10	US-09-864-761-26473	Sequence 26473, A
9	12.6	63.0	96	10	US-09-864-761-28720	Sequence 28720, A
10	12.4	62.0	30	10	US-09-875-519A-2	Sequence 2, Appli
c 11	12.4	62.0	30	10	US-09-875-519A-5	Sequence 5, Appli
c 12	12.4	62.0	31	10	US-09-801-274-1274	Sequence 1274, Ap
13	12	60.0	24	10	US-09-823-936-8	Sequence 8, Appli
14	12	60.0	24	10	US-09-755-836-7	Sequence 7, Appli
c 15	12	60.0	25	10	US-09-877-095-3	Sequence 3, Appli
16	12	60.0	60	10	US-09-886-900-6	Sequence 6, Appli
c 17	12	60.0	87	10	US-09-864-761-22774	Sequence 22774, A
18	12	60.0	100	10	US-09-851-190A-6	Sequence 6, Appli
19	11.8	59.0	31	10	US-09-801-274-1422	Sequence 1422, Ap
20	11.8	59.0	35	9	US-10-023-096-6	Sequence 6, Appli
21	11.8	59.0	46	10	US-09-920-171-28	Sequence 28, Appl
22	11.8	59.0	78	10	US-09-864-761-28519	Sequence 28519, A
23	11.8	59.0	84	9	US-10-109-971-7	Sequence 7, Appli
24	11.8	59.0	88	10	US-09-864-761-30659	Sequence 30659, A
25	11.8	59.0	95	10	US-09-864-761-23346	Sequence 23346, A
26	11.6	58.0	30	10	US-09-875-519A-8	Sequence 8, Appli
c 27	11.6	58.0	31	9	US-10-109-971-2	Sequence 2, Appli
c 28	11.6	58.0	31	9	US-10-109-971-8	Sequence 8, Appli
c 29	11.6	58.0	31	10	US-09-801-274-1045	Sequence 1045, Ap
30	11.6	58.0	72	9	US-09-916-494A-17	Sequence 17, Appl
31	11.6	58.0	75	10	US-09-880-253A-5	Sequence 5, Appli
32	11.6	58.0	75	10	US-09-880-253A-56	Sequence 56, Appl
33	11.6	58.0	75	10	US-09-880-253A-57	Sequence 57, Appl
34	11.6	58.0	75	10	US-09-880-253A-58	Sequence 58, Appl
c 35	11.6	58.0	80	10	US-09-864-761-28043	Sequence 28043, A
36	11.4	57.0	17	10	US-09-866-108-9221	Sequence 9221, Ap
37	11.4	57.0	17	10	US-09-866-108-9222	Sequence 9222, Ap
38	11.4	57.0	17	10	US-09-866-108-9223	Sequence 9223, Ap
39	11.4	57.0	17	10	US-09-866-108-9224	Sequence 9224, Ap
40	11.4	57.0	17	10	US-09-866-108-9225	Sequence 9225, Ap
41	11.4	57.0	25	10	US-09-866-108-14113	Sequence 14113, A
42	11.4	57.0	25	10	US-09-866-108-14114	Sequence 14114, A
43	11.4	57.0	25	10	US-09-866-108-14115	Sequence 14115, A
44	11.4	57.0	25	10	US-09-866-108-14116	Sequence 14116, A
45	11.4	57.0	25	10	US-09-866-108-14117	Sequence 14117, A

ALIGNMENTS

RESULT 1
US-09-791-378-664
; Sequence 664, Application US/09791378
; Patent No. US20020142303A1
; GENERAL INFORMATION:
; APPLICANT: Parekh, Rajesh
; TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: SCHIZOPHRENIA
; FILE REFERENCE: 9195-061-999
; CURRENT APPLICATION NUMBER: US/09/791,378
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/750,395
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 677
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 664
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-791-378-664

Query Match 71.0%; Score 14.2; DB 10; Length 42;
Best Local Similarity 84.2%; Pred. No. 3.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 GTGCCCAGCCAGAGCGACT 19
          ||| ||| | |||||||||
Db      7 GTGGCCATCGAGAGCGACT 25
```

RESULT 2
US-09-864-761-18482/c
; Sequence 18482, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1

```
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18482
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL117694.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =   4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL =  14
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.8
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
;   OTHER INFORMATION: SWISSPROT HIT: P12814, EVALUE 8.00e-12
;   OTHER INFORMATION: NT HIT: gi11435217, EVALUE 1.00e-49
;   OTHER INFORMATION: EST_HUMAN HIT: BE831655.1, EVALUE 1.00e-49
US-09-864-761-18482

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 100;
  Best Local Similarity   93.3%;  Pred. No. 8.6e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 GCCCAGCCAGAGCGA 17
           ||| |||||||||||
Db      89 GCCAAGCCAGAGCGA 75


RESULT 3
US-09-559-671A-54
; Sequence 54, Application US/09559671A
; Patent No. US20020051976A1
; GENERAL INFORMATION:
;  APPLICANT: Patten, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYPEPTIDE ENGINEERING
;  FILE REFERENCE: 02-020503US
;  CURRENT APPLICATION NUMBER: US/09/559,671A
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: 08/769,062
;  PRIOR FILING DATE: 1996-12-18
;  PRIOR APPLICATION NUMBER: 08/198,431
;  PRIOR FILING DATE: 1994-02-17
;  PRIOR APPLICATION NUMBER: 08/425,684
;  PRIOR FILING DATE: 1995-04-18
;  PRIOR APPLICATION NUMBER: 08/537,874
;  PRIOR FILING DATE: 1995-10-30
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
;   LENGTH: 58
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: degenerate
;   OTHER INFORMATION: oligonucleotide used for codon usage library
US-09-559-671A-54

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 58;
  Best Local Similarity   81.2%;  Pred. No. 1.6e+03;
  Matches   13;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGACTG 20
           ||||||||:|| ||:|
Db      15 CCAGCCAGHGCHACHG 30


RESULT 4
US-09-760-008A-4/c
; Sequence 4, Application US/09760008A
; Patent No. US20020004483A1
; GENERAL INFORMATION:
;  APPLICANT: NISSEN, TORBEN LAUESGAARD
;  APPLICANT:  ANDERSEN, KIM VILBOUR
;  APPLICANT:  HANSEN, CHRISTIAN KARSTEN
;  APPLICANT:  MIKKELSEN, JAN MOLLER
;  TITLE OF INVENTION: G-CSF CONJUGATES
;  FILE REFERENCE: 31-000700US
;  CURRENT APPLICATION NUMBER: US/09/760,008A
;  CURRENT FILING DATE:  2001-05-11
;  PRIOR APPLICATION NUMBER: 60/176,376
;  PRIOR FILING DATE: 2000-01-14
;  PRIOR APPLICATION NUMBER: 60/189,506
;  PRIOR FILING DATE: 2000-03-15
;  PRIOR APPLICATION NUMBER: 60/215,644
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DK PA 2000 00024
;  PRIOR FILING DATE: 2000-01-10
;  PRIOR APPLICATION NUMBER: DK PA 2000 00341
;  PRIOR FILING DATE: 2000-03-02
;  PRIOR APPLICATION NUMBER: DK PA 2000 00943
;  PRIOR FILING DATE: 2000-06-16
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 63
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-09-760-008A-4

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 63;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+03;
```

```
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps  0;

Qy      5 CCAGCCAGAGCGACTG 20
          |||||||| || ||||
Db     41 CCAGCCAGTGCCACTG 26


RESULT 5
US-09-875-494-20/c
; Sequence 20, Application US/09875494
; Patent No. US20020072093A1
; GENERAL INFORMATION:
;  APPLICANT: Schmidt, Brian
;  APPLICANT:  Chen, Tseng-hui
;  TITLE OF INVENTION: No. US20020072093A1el Leader Peptides for Enhancing Secretion of
;  TITLE OF INVENTION:  a Host Cell
;  FILE REFERENCE: COUL-012-01
;  CURRENT APPLICATION NUMBER: US/09/875,494
;  CURRENT FILING DATE:  2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/209,517
;  PRIOR FILING DATE: 2000-06-05
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-09-875-494-20

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 65;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 CCAGCCAGAGCGACTG 20
          |||||||| || ||||
Db     43 CCAGCCAGTGCCACTG 28


RESULT 6
US-10-004-832-3/c
; Sequence 3, Application US/10004832
; Patent No. US20020146408A1
; GENERAL INFORMATION:
;  APPLICANT: Kirkness et al.
;  TITLE OF INVENTION: Human Haemopoietic Maturation Factor
;  FILE REFERENCE: PF105P1D2
;  CURRENT APPLICATION NUMBER: US/10/004,832
;  CURRENT FILING DATE:  2001-12-07
;  PRIOR APPLICATION NUMBER: US 09/333,033
;  PRIOR FILING DATE: 1999-06-15
;  PRIOR APPLICATION NUMBER: US 08/442,497
;  PRIOR FILING DATE: 1995-05-16
;  PRIOR APPLICATION NUMBER: US 08/187,186
;  PRIOR FILING DATE: 1994-01-25
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 93
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Contains a BspHI restriction enzyme site and the ompA leader sequ
;   OTHER INFORMATION: ence.
US-10-004-832-3

  Query Match             64.0%;  Score 12.8;  DB 12;  Length 93;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 CCAGCCAGAGCGACTG 20
          |||||||| || ||||
Db     47 CCAGCCAGTGCCACTG 32


RESULT 7
US-09-815-343-1484/c
; Sequence 1484, Application US/09815343
; Patent No. US20010055596A1
; GENERAL INFORMATION:
;  APPLICANT: Meagher, Madeleine
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.504
;  CURRENT APPLICATION NUMBER: US/09/815,343
;  CURRENT FILING DATE:  2001-03-22
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1484
;   LENGTH: 99
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-815-343-1484

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 99;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 TGCCCAGCCAGAGCGA 17
          ||||| || |||||||
Db     35 TGCCCTGCTAGAGCGA 20


RESULT 8
US-09-864-761-26473/c
; Sequence 26473, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26473
;   LENGTH: 79
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC006548.19
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
;   OTHER INFORMATION: NT HIT: X83528.1, EVALUE 6.30e-01
;   OTHER INFORMATION: EST_HUMAN HIT: AI126906.1, EVALUE 9.40e-01
;   OTHER INFORMATION: SWISSPROT HIT: Q99759, EVALUE 5.60e+00
US-09-864-761-26473

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 79;
  Best Local Similarity  78.9%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 TGCCCAGCCAGAGCGACTG 20
          ||  |||||||||  | |||
Db     31 TGGGCAGCCAGAGAGCCTG 13


RESULT 9
US-09-864-761-28720
; Sequence 28720, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28720
;   LENGTH: 96
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC013597.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.98
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
;   OTHER INFORMATION: NT HIT: Y11095.1, EVALUE 8.00e-01
US-09-864-761-28720

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 96;
  Best Local Similarity  78.9%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 TGCCCAGCCAGAGCGACTG 20
          |||| || |||||  ||||
Db     51 TGCCTAGTCAGAGGCACTG 69


RESULT 10
US-09-875-519A-2
; Sequence 2, Application US/09875519A
; Patent No. US20020068059A1
; GENERAL INFORMATION:
;  APPLICANT: Farries, Timothy C.
;  APPLICANT:  Harrison, Richard A.
;  TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
;  FILE REFERENCE: 4-30443/A/IMU/PCT
;  CURRENT APPLICATION NUMBER: US/09/875,519A
;  CURRENT FILING DATE:  2001-06-06
;  PRIOR APPLICATION NUMBER: PCT/GB97/00603
;  PRIOR FILING DATE: 1997-03-04
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  mutagenic
;   OTHER INFORMATION: oligodeoxynucleotide QRI1
US-09-875-519A-2

  Query Match            62.0%;  Score 12.4;  DB 10;  Length 30;
  Best Local Similarity  92.9%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TGCCCAGCCAGAGC 15
          |||||||||| |||
Db      5 TGCCCAGCCAAAGC 18
```

```
RESULT 11
US-09-875-519A-5/c
; Sequence 5, Application US/09875519A
; Patent No. US20020068059A1
; GENERAL INFORMATION:
;  APPLICANT: Farries, Timothy C.
;  APPLICANT:  Harrison, Richard A.
;  TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
;  FILE REFERENCE: 4-30443/A/IMU/PCT
;  CURRENT APPLICATION NUMBER: US/09/875,519A
;  CURRENT FILING DATE:  2001-06-06
;  PRIOR APPLICATION NUMBER: PCT/GB97/00603
;  PRIOR FILING DATE: 1997-03-04
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  antisense
;   OTHER INFORMATION: oligodeoxynucleotide QRI1n
US-09-875-519A-5

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 30;
  Best Local Similarity   92.9%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 TGCCCAGCCAGAGC 15
           |||||||||| |||
Db      26 TGCCCAGCCAAAGC 13


RESULT 12
US-09-801-274-1274/c
; Sequence 1274, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1274
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-1274

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 31;
  Best Local Similarity   92.9%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 AGCCAGAGCGACTG 20
           |||||||| |||||
Db      30 AGCCAGAGGGACTG 17


RESULT 13
US-09-823-936-8
; Sequence 8, Application US/09823936
; Patent No. US20020061309A1
; GENERAL INFORMATION:
;  APPLICANT: GARGER, Stephen J.
;  APPLICANT:  GROSS,  Cynthia
;  APPLICANT:  LINDBO, John A.
;  APPLICANT:  POGUE, Gregory P.
;  TITLE OF INVENTION: PRODUCTION OF PEPTIDES IN PLANTS AS
;  TITLE OF INVENTION:  N-TERMINAL VIRAL COAT PROTEIN FUSIONS
;  FILE REFERENCE: 008010087CPUS05
;  CURRENT APPLICATION NUMBER: US/09/823,936
;  CURRENT FILING DATE:  2001-03-29
;  PRIOR APPLICATION NUMBER: 09/ 520,967
;  PRIOR FILING DATE: 2000-03-08
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Tobacco mosaic virus tobamovirus
US-09-823-936-8

  Query Match             60.0%;  Score 12;  DB 10;  Length 24;
  Best Local Similarity   75.0%;  Pred. No. 3.5e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 GTGCCCAGCCAGAGCGACTG 20
           ||  |||||| || || |||
Db       3 GTCACCAGCCCGATCGCCTG 22


RESULT 14
US-09-755-836-7
; Sequence 7, Application US/09755836
; Patent No. US20020107387A1
;  GENERAL INFORMATION:
;       APPLICANT: Turpen, Thomas H.
;                  Reinl, Stephen
;                  Grill, Laurence K.
;       TITLE OF INVENTION: Production of Peptides in Plants as
;                           Viral Coat Protein Fusions
;       NUMBER OF SEQUENCES: 27
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Pennie & Edmonds
;            STREET: 1155 Avenue of the Americas
;            CITY: New York
;            STATE: New York
;            COUNTRY: USA
;            ZIP: 10036
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/755,836
;            FILING DATE: 05-Jan-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/324,003
;            FILING DATE: 14-OCT-1994
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Halluin, Albert P.
;            REGISTRATION NUMBER: 25,227
;            REFERENCE/DOCKET NUMBER: 8129-087
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 415-854-3660
;            TELEFAX: 415-854-3694
;            TELEX: 66141 PENNIE
;  INFORMATION FOR SEQ ID NO: 7:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 24 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: unknown
;            TOPOLOGY: unknown
;       MOLECULE TYPE: DNA (genomic)
```

```
;       SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-755-836-7

  Query Match             60.0%;  Score 12;  DB 10;  Length 24;
  Best Local Similarity   75.0%;  Pred. No. 3.5e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 GTGCCCAGCCAGAGCGACTG 20
           ||  |||||| || || |||
Db       3 GTCACCAGCCCGATCGCCTG 22


RESULT 15
US-09-877-095-3/c
; Sequence 3, Application US/09877095
; Patent No. US20020123051A1
; GENERAL INFORMATION:
;  APPLICANT: Danenberg, K. et al.
;  TITLE OF INVENTION: METHOD OF DETERMINING A CHEMOTHERAPEUTIC
;  TITLE OF INVENTION:  REGIMEN  BASED ON ERCC1 EXPRESSION
;  FILE REFERENCE: 11220/127
;  CURRENT APPLICATION NUMBER: US/09/877,095
;  CURRENT FILING DATE:  2001-06-11
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide Primer
US-09-877-095-3

  Query Match             60.0%;  Score 12;  DB 10;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 3.5e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 GTGCCCAGCCAGAGCGACTG 20
           ||||   ||||||||  |||
Db      22 GTGCTGGGCCAGAGCACCTG 3


Search completed: November 23, 2002, 06:42:07
Job time : 18.25 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-8
Perfect score: 20
Sequence: 1 gtgcccagccagagcgactg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	15.8	79.0	64	17	AZ783365	AZ783365 2M0025F03
c	2	15.2	76.0	85	9	AA797827	AA797827 vy04b03.r
c	3	13.8	69.0	64	9	AI037190	AI037190 ud60a09.y
	4	13.6	68.0	55	17	AZ487430	AZ487430 1M0317A10
c	5	13.4	67.0	20	17	AZ946089	AZ946089 2M0207A13
	6	13.4	67.0	86	17	AZ475290	AZ475290 1M0293H07
c	7	13.4	67.0	95	14	T80645	T80645 yd92f03.s1
c	8	13.2	66.0	50	9	AU104251	AU104251 AU104251
	9	13.2	66.0	50	14	H74150	H74150 ys16c04.r1
c	10	13.2	66.0	58	17	BH635408	BH635408 1008004F0
	11	13.2	66.0	73	9	AI158821	AI158821 ud38e10.r
	12	13	65.0	76	9	AA118149	AA118149 mn11a12.r
c	13	13	65.0	82	13	BI902359	BI902359 ib86d03.x
c	14	12.8	64.0	42	17	AZ811797	AZ811797 2M0078P06
c	15	12.8	64.0	50	9	AU102538	AU102538 AU102538
c	16	12.8	64.0	51	9	AI833237	AI833237 at76d05.x
	17	12.8	64.0	52	10	AW059751	AW059751 LE3c07.yg
	18	12.8	64.0	76	9	AA451089	AA451089 vf87d06.r
	19	12.8	64.0	79	9	AI019859	AI019859 ua92h04.r
c	20	12.8	64.0	83	9	AA587260	AA587260 nn70f05.s
	21	12.8	64.0	85	9	AA486861	AA486861 ab16g07.r
	22	12.8	64.0	94	9	AA571972	AA571972 vh18a03.r
c	23	12.8	64.0	97	9	AA837701	AA837701 oe06c02.s
	24	12.8	64.0	98	14	W99199	W99199 mf59f12.r1
	25	12.8	64.0	100	10	AW601761	AW601761 MR4-BT040
c	26	12.8	64.0	100	10	AW903646	AW903646 CM4-NN103
c	27	12.6	63.0	23	17	AZ605727	AZ605727 1M0427A15
	28	12.6	63.0	34	17	AZ826457	AZ826457 2M0102P11
	29	12.6	63.0	48	17	BH808776	BH808776 1008082D0
	30	12.6	63.0	52	9	AI477512	AI477512 fb57c08.x
c	31	12.6	63.0	62	17	BH849932	BH849932 SALK_0705
c	32	12.6	63.0	65	17	AZ494609	AZ494609 1M0330B03
c	33	12.6	63.0	70	17	AZ389571	AZ389571 1M0150E24
	34	12.6	63.0	82	9	AA664777	AA664777 nu68f02.s
	35	12.6	63.0	85	9	AI905106	AI905106 QV-BT077-
	36	12.6	63.0	88	9	AA926857	AA926857 om07f09.s
	37	12.6	63.0	89	9	AI905184	AI905184 QV-BT077-
c	38	12.6	63.0	89	9	AI905194	AI905194 QV-BT077-
	39	12.6	63.0	91	9	AI738408	AI738408 wi23e05.x
c	40	12.6	63.0	92	9	AI932967	AI932967 wo40f09.x
	41	12.6	63.0	93	17	BH229029	BH229029 1006149G0
c	42	12.6	63.0	94	9	AI935190	AI935190 wp14h06.x
	43	12.6	63.0	100	12	BF953180	BF953180 CM3-NN118
c	44	12.4	62.0	37	17	BH840631	BH840631 KG05933-5
	45	12.4	62.0	42	17	BH621959	BH621959 1007115G1

ALIGNMENTS

RESULT 1
AZ783365
LOCUS AZ783365 64 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0025F03F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0025F03 F, DNA sequence.
ACCESSION AZ783365
VERSION AZ783365.1 GI:12918022
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 64)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

```
            Plate: 0025  row: F  column: 03
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 64.
FEATURES              Location/Qualifiers
     source           1. .64
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC2M0025F03"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT       14 a     18 c     13 g     19 t
ORIGIN

  Query Match            79.0%;  Score 15.8;  DB 17;  Length 64;
  Best Local Similarity  89.5%;  Pred. No. 2.6e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 TGCCCAGCCAGAGCGACTG 20
           |||||| ||||||| ||||
Db      31 TGCCCACCCAGAGCAACTG 49


RESULT 2
AA797827/c
LOCUS       AA797827                 85 bp    mRNA    linear   EST 10-FEB-1998
DEFINITION  vy04b03.r1 Stratagene mouse macrophage (#937306) Mus musculus cDNA
            clone IMAGE:1294445 5' similar to SW:YMN0_YEAST Q03103 HYPOTHETICAL
            65.0 KD PROTEIN IN COX14 5'REGION PRECURSOR. ;, mRNA sequence.
ACCESSION   AA797827
VERSION     AA797827.1  GI:2860782
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:675493
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .85
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:1294445"
                      /clone_lib="Stratagene mouse macrophage (#937306)"
                      /tissue_type="macrophage"
                      /dev_stage="WEHI-3 cell line"
                      /lab_host="SOLR (kanamycin resistant)"
                      /note="Organ: blood; Vector: pBluescript SK-; Site_1:
                      EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                      Oligo dT. WEHI-3 cell line. Average insert size: 1.5 kb;
                      Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                      3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       13 a     21 c     28 g     23 t
ORIGIN

  Query Match            76.0%;  Score 15.2;  DB 9;  Length 85;
  Best Local Similarity  85.0%;  Pred. No. 5.3e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCCCAGCCAGAGCGACTG 20
           |||||||||| | ||| |||
Db      66 GTGCCCAGCCCGTGCGTCTG 47


RESULT 3
AI037190/c
LOCUS       AI037190                 64 bp    mRNA    linear   EST 26-JUN-1998
DEFINITION  ud60a09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1450264 5' similar to SW:YHL1_EBV P03181 HYPOTHETICAL BHLF1
            PROTEIN. ;, mRNA sequence.
ACCESSION   AI037190
VERSION     AI037190.1  GI:3260881
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:923580
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: custom primer used
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .64
                      /organism="Mus musculus"
                      /strain="C57BL"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:1450264"
```

/clone_lib="Sugano mouse liver mlia"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."
BASE COUNT 0 a 26 c 24 g 14 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 64;
Best Local Similarity 88.2%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 GTGCCCAGCCAGAGCGA 17
          | | ||||||||||||||
Db     44 GAGACCAGCCAGAGCGA 28
```

RESULT 4
AZ487430
LOCUS AZ487430 55 bp DNA linear GSS 05-OCT-2000
DEFINITION 1M0317A10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0317A10 F, DNA sequence.
ACCESSION AZ487430
VERSION AZ487430.1 GI:10655167
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 55)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0317 row: A column: 10
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 55.
FEATURES Location/Qualifiers
source 1. .55
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0317A10"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 15 a 17 c 19 g 4 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 55;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 GTGCCCAGCCAGAGCGACTG 20
          | |||||||||| | || ||
Db     30 GAGCCCAGCCAGGGTGAGTG 49
```

RESULT 5
AZ946089/c
LOCUS AZ946089 20 bp DNA linear GSS 27-APR-2001
DEFINITION 2M0207A13R Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0207A13 R, DNA sequence.
ACCESSION AZ946089
VERSION AZ946089.1 GI:13812898
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0207 row: A column: 13
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 20.
FEATURES Location/Qualifiers
source 1. .20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0207A13"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson

Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 4 a 4 c 7 g 5 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 20;
Best Local Similarity 93.3%; Pred. No. 2e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CCAGCCAGAGCGACT 19
||||||||||||| |
Db 18 CCAGCCAGAGCGATT 4

RESULT 6
AZ475290
LOCUS AZ475290 86 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0293H07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0293H07 F, DNA sequence.
ACCESSION AZ475290
VERSION AZ475290.1 GI:10633415
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 86)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0293 row: H column: 07
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 86.
FEATURES Location/Qualifiers
source 1. .86
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0293H07"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 17 a 27 c 22 g 20 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 86;
Best Local Similarity 93.3%; Pred. No. 3.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTGCCCAGCCAGAGC 15
|||||||||||| ||
Db 34 GTGCCCAGCCAGTGC 48

RESULT 7
T80645/c
LOCUS T80645 95 bp mRNA linear EST 15-MAR-1995
DEFINITION yd92f03.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:115709 3' similar to gb:J04970 CARBOXYPEPTIDASE M PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION T80645
VERSION T80645.1 GI:703530
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 95)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 727
High quality sequence starts: 1 High quality sequence stops: 1
Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality
Insert Length: 727 Std Error: 0.00
Seq primer: -21m13
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .95
/organism="Homo sapiens"
/db_xref="GDB:471326"
/db_xref="taxon:9606"
/clone="IMAGE:115709"
/clone_lib="Soares fetal liver spleen 1NFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"

/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 18 a 22 c 24 g 25 t 6 others
ORIGIN

Query Match 67.0%; Score 13.4; DB 14; Length 95;
Best Local Similarity 87.5%; Pred. No. 3.5e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GCCCAGCCAGAGCGAC 18
| |||||||||||
Db 35 GNCCAGCCAGAGCAAC 20

RESULT 8
AU104251/c
LOCUS AU104251 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU104251 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HEP18556, mRNA sequence.
ACCESSION AU104251
VERSION AU104251.1 GI:13553772
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP18556"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and dimethylfumarate treated U937 cells"
BASE COUNT 10 a 14 c 15 g 11 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 50;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GCCCAGCCAGAGCGACTG 20
|| ||||||||| ||||
Db 22 GCGCAGCCAGAGTTACTG 5

RESULT 9
H74150
LOCUS H74150 50 bp mRNA linear EST 31-OCT-1995
DEFINITION ys16c04.r1 Soares breast 2NbHBst Homo sapiens cDNA clone IMAGE:214950 5' similar to gb:J05481 ENDOGLIN PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION H74150
VERSION H74150.1 GI:1047412
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 3344
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 3344 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:214950"
/clone_lib="Soares breast 2NbHBst"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: breast; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 230. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 8 a 20 c 18 g 3 t 1 others
ORIGIN

Query Match 66.0%; Score 13.2; DB 14; Length 50;
Best Local Similarity 83.3%; Pred. No. 3.4e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GCCCAGCCAGAGCGACTG 20
||||||| ||| |||||
Db 16 GCCCAGCAGGAGAGACTG 33

RESULT 10
BH635408/c
LOCUS BH635408 58 bp DNA linear GSS 14-FEB-2002
DEFINITION 1008004F07.2EL_x1 1008 - RescueMu Grid I Zea mays genomic, DNA sequence.
ACCESSION BH635408
VERSION BH635408.1 GI:18657645

```
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1008004  row: 12
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1008 - RescueMu Grid I"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the  web
                     site www.zmdb.iastate.edu and follow the links for
                     'RescueMu.' Grid I was grown at Berkeley in 2001. DNA was
                     extracted from leaf punches,  double digested using BamHI
                     and BglII, and ligated to form circular plasmids.  DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       11 a     11 c     19 g     17 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 58;
  Best Local Similarity   83.3%;  Pred. No. 3.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCCCAGCCAGAGCGACTG 20
           || |||||  ||||||||
Db      32 GCTCAGCCTCAGCGACTG 15


RESULT 11
AI158821
LOCUS       AI158821                  73 bp    mRNA    linear   EST 02-OCT-1998
DEFINITION  ud38e10.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
            IMAGE:1448202 5' similar to SW:RM12_CRICR P52827 MITOCHONDRIAL 60S
            RIBOSOMAL PROTEIN L7/L12 PRECURSOR ;, mRNA sequence.
ACCESSION   AI158821
VERSION     AI158821.1  GI:3692003
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:921518
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28ml3 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1448202"
                     /clone_lib="Soares_mammary_gland_NbMMG"
                     /sex="male"
                     /tissue_type="mammary gland"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of  the modified pT7T3 vector.
                     RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
BASE COUNT       11 a     24 c     27 g     11 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 73;
  Best Local Similarity   83.3%;  Pred. No. 3.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCCCAGCCAGAGCGACTG 20
           ||| |||||||||||| |
Db      33 GCCTTGCCAGAGCGACAG 50


RESULT 12
AA118149
LOCUS       AA118149                  76 bp    mRNA    linear   EST 15-NOV-1996
DEFINITION  mn11a12.r1 Beddington mouse embryonic region Mus musculus cDNA
            clone IMAGE:537598 5' similar to gb:X03559 ATP SYNTHASE BETA CHAIN,
            MITOCHONDRIAL PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AA118149
VERSION     AA118149.1  GI:1673180
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
```

```
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:324534
            Putative full length read
            vector to vector length is 172
            Seq primer: -40m13 ET.
FEATURES              Location/Qualifiers
     source           1. .76
                      /organism="Mus musculus"
                      /strain="C57BL/6 x DBA"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:537598"
                      /clone_lib="Beddington mouse embryonic region"
                      /sex="pooled"
                      /tissue_type="embryo"
                      /dev_stage="7.5dpc"
                      /lab_host="DH12S"
                      /note="Organ: whole embryo; Vector: pCMV-SPORT; Site_1:
                      SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                      Oligo dT. Gastrulating embryos were collected at 7.5dpc
                      from C57BL6 x DBA matings, excluding embryos that had
                      developed head folds and all extraembryonic tissues.
                      Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
                      Referenced in Development 121, 2479-2489 (1995)"
BASE COUNT       20 a     26 c     16 g     14 t
ORIGIN

  Query Match             65.0%;  Score 13;  DB 9;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 CCAGCCAGAGCGA 17
           |||||||||||||
Db       1 CCAGCCAGAGCGA 13


RESULT 13
BI902359/c
LOCUS       BI902359                 82 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  ib86d03.x4 Amplified Melton Mouse Islets 1 MIS1-A Mus musculus cDNA
            clone IMAGE:5653660 3' similar to SW:TRYP_MOUSE P07146 TRYPSINOGEN
            PRECURSOR ;, mRNA sequence.
ACCESSION   BI902359
VERSION     BI902359.1  GI:16195485
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.
            , Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ib86d03.y4
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Douglas Melton DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Juliana Brown
            (brown@fas.harvard.edu)
            Trace considered overall poor quality
            MGI:1939986 This sequence now available from the IMAGE consortium,
            for clone orders contact: info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .82
                      /organism="Mus musculus"
                      /strain="ICR"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:5653660"
                      /clone_lib="Amplified Melton Mouse Islets 1 MIS1-A"
                      /sex="Male"
                      /tissue_type="Islets of Langerhans"
                      /dev_stage="Adult"
                      /lab_host="DH10B"
                      /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
                      Site_2: Sal 1; Library constructed using SuperScript
                      Plasmid Library kit (Life  Technologies). cDNA made by
                      oligo-dT priming. Size-selected by column  fractionation;
                      average insert size 0.91 kb. Amplified once on solid
                      support. cDNA Library Preparation: Guolin Chen. "
BASE COUNT       18 a     21 c     22 g     21 t
ORIGIN

  Query Match             65.0%;  Score 13;  DB 13;  Length 82;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GTGCCCAGCCAGA 13
           |||||||||||||
Db      78 GTGCCCAGCCAGA 66


RESULT 14
AZ811797/c
LOCUS       AZ811797                 42 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0078P06F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0078P06 F, DNA sequence.
ACCESSION   AZ811797
VERSION     AZ811797.1  GI:12980410
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0078  row: P  column: 06
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 42.
FEATURES              Location/Qualifiers
     source           1. .42
                      /organism="Mus musculus"
```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0078P06"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 8 a 8 c 15 g 11 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 17; Length 42;
Best Local Similarity 87.5%; Pred. No. 4.9e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CCCAGCCAGAGCGACT 19
| |||||||||||| |
Db 29 CTCAGCCAGAGCGATT 14

RESULT 15
AU102538/c
LOCUS AU102538 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU102538 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HRC03012, mRNA sequence.
ACCESSION AU102538
VERSION AU102538.1 GI:13552059
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HRC03012"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and dimethylfumarate treated U937 cells"

BASE COUNT 6 a 11 c 23 g 10 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 9; Length 50;
Best Local Similarity 87.5%; Pred. No. 5.2e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CCCAGCCAGAGCGACT 19
||| ||||||| ||||
Db 16 CCCCGCCAGAGGGACT 1

Search completed: November 26, 2002, 04:07:37
Job time : 767.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-9
Perfect score: 20
Sequence: 1 tgaggtgcgggtggaagtgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	15	75.0	24	6	A57512	A57512 Sequence 4
c 2	15	75.0	24	6	A57514	A57514 Sequence 6
c 3	15	75.0	24	6	AR052978	AR052978 Sequence
c 4	15	75.0	24	6	AR052980	AR052980 Sequence
5	14.8	74.0	48	6	A28016	A28016 pCD4-FvCD3
c 6	14.8	74.0	48	6	A28017	A28017 pCD4-FvCD3
7	14.4	72.0	100	9	AJ8144	AJ008144 Homo sapi
c 8	14.2	71.0	30	6	AR106307	AR106307 Sequence
c 9	14.2	71.0	30	6	AR110116	AR110116 Sequence
c 10	14.2	71.0	30	6	AR126277	AR126277 Sequence
11	14.2	71.0	51	6	AX161940	AX161940 Sequence
c 12	14.2	71.0	85	3	AF318496	AF318496 Scutigere
c 13	14	70.0	24	6	A57510	A57510 Sequence 2
c 14	14	70.0	24	6	A57516	A57516 Sequence 8
c 15	14	70.0	24	6	A57518	A57518 Sequence 10
c 16	14	70.0	24	6	AR052976	AR052976 Sequence
c 17	14	70.0	24	6	AR052982	AR052982 Sequence
c 18	14	70.0	24	6	AR052984	AR052984 Sequence
c 19	13.8	69.0	24	6	AX353651	AX353651 Sequence
20	13.8	69.0	38	6	AX247606	AX247606 Sequence
c 21	13.8	69.0	48	6	A28937	A28937 Oligonucleo
c 22	13.8	69.0	48	6	I12699	I12699 Sequence 20
23	13.8	69.0	97	10	AF177699	AF177699 Rattus no
24	13.6	68.0	23	6	AX394046	AX394046 Sequence
25	13.6	68.0	57	6	A09304	A09304 Serratia sp
c 26	13.6	68.0	57	6	A09305	A09305 Serratia sp
c 27	13.6	68.0	96	6	A74696	A74696 Sequence 38
c 28	13.6	68.0	96	6	A77675	A77675 Sequence 38
c 29	13.4	67.0	27	6	A86669	A86669 Sequence 15
c 30	13.4	67.0	39	12	SYNANVAJ	M60038 Avian neovi
31	13.4	67.0	97	6	I40748	I40748 Sequence 79
c 32	13.2	66.0	25	6	AX174830	AX174830 Sequence
33	13.2	66.0	30	6	AR177831	AR177831 Sequence
c 34	13.2	66.0	31	6	AX248112	AX248112 Sequence
35	13.2	66.0	51	6	AR022080	AR022080 Sequence
c 36	13.2	66.0	51	6	AR022081	AR022081 Sequence
37	13.2	66.0	51	6	I92870	I92870 Sequence 44
c 38	13.2	66.0	51	6	I92871	I92871 Sequence 45
c 39	13.2	66.0	57	5	GGTHR5A2	Y13977 Gallus gall
40	13.2	66.0	65	5	GGIGY1	AF190139 Gallus ga
c 41	13.2	66.0	65	6	AX485704	AX485704 Sequence
c 42	13.2	66.0	90	6	AX241114	AX241114 Sequence
43	12.8	64.0	33	6	AX395702	AX395702 Sequence
44	12.8	64.0	65	6	AX484981	AX484981 Sequence
45	12.8	64.0	69	6	AR092211	AR092211 Sequence

ALIGNMENTS

RESULT 1
A57512/c
LOCUS A57512 24 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 4 from Patent WO9632483.
ACCESSION A57512
VERSION A57512.1 GI:3713370
KEYWORDS .
SOURCE unidentified.
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 24)
AUTHORS Masucci,M.G.
TITLE IMMUNE-EVADING PROTEINS
JOURNAL Patent: WO 9632483-A 4 17-OCT-1996;
MASUCCI MARIA GRAZIA (SE)

```
COMMENT     Other publication AU 5284296 961030.
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        4 a     14 c      2 g      4 t
ORIGIN

  Query Match             75.0%;  Score 15;  DB 6;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAA 16
          |||||||||||||||
Db     15 GAGGTGCGGGTGGAA 1


RESULT 2
A57514/c
LOCUS       A57514                    24 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 6 from Patent WO9632483.
ACCESSION   A57514
VERSION     A57514.1  GI:3713372
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Masucci,M.G.
  TITLE     IMMUNE-EVADING PROTEINS
  JOURNAL   Patent: WO 9632483-A 6 17-OCT-1996;
            MASUCCI MARIA GRAZIA (SE)
COMMENT     Other publication AU 5284296 961030.
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        3 a     14 c      2 g      5 t
ORIGIN

  Query Match             75.0%;  Score 15;  DB 6;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAA 16
          |||||||||||||||
Db     15 GAGGTGCGGGTGGAA 1


RESULT 3
AR052978/c
LOCUS       AR052978                  24 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 7 from patent US 5833991.
ACCESSION   AR052978
VERSION     AR052978.1  GI:5977840
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Masucci,M.G.
  TITLE     Glycine-containing sequences conferring invisibility to the immune
            system
  JOURNAL   Patent: US 5833991-A 7 10-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        4 a     14 c      2 g      4 t
ORIGIN

  Query Match             75.0%;  Score 15;  DB 6;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAA 16
          |||||||||||||||
Db     15 GAGGTGCGGGTGGAA 1


RESULT 4
AR052980/c
LOCUS       AR052980                  24 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 10 from patent US 5833991.
ACCESSION   AR052980
VERSION     AR052980.1  GI:5977842
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Masucci,M.G.
  TITLE     Glycine-containing sequences conferring invisibility to the immune
            system
  JOURNAL   Patent: US 5833991-A 10 10-NOV-1998;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        3 a     14 c      2 g      5 t
ORIGIN

  Query Match             75.0%;  Score 15;  DB 6;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAA 16
          |||||||||||||||
Db     15 GAGGTGCGGGTGGAA 1


RESULT 5
A28016
LOCUS       A28016                    48 bp    DNA     linear   PAT 03-OCT-1995
DEFINITION  pCD4-FvCD3 construction oligo 2119.
ACCESSION   A28016
VERSION     A28016.1  GI:1248563
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Karjalainen,K., Lanzavecchia,A. and Traunecker,A.
  TITLE     Chimaeric polypeptides
  JOURNAL   Patent: EP 0505908-A 11 30-SEP-1992;
            F. HOFFMANN-LA ROCHE AG
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       10 a      3 c     24 g     11 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 6;  Length 48;
  Best Local Similarity   88.9%;  Pred. No. 2.6e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAAGTG 19
          ||||| | ||||||||||
Db     16 GAGGTTCTGGTGGAAGTG 33


RESULT 6
A28017/c
LOCUS       A28017                    48 bp    DNA     linear   PAT 03-OCT-1995
DEFINITION  pCD4-FvCD3 construction oligo 2120.
ACCESSION   A28017
VERSION     A28017.1  GI:1248564
```

```
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Karjalainen,K., Lanzavecchia,A. and Traunecker,A.
  TITLE     Chimaeric polypeptides
  JOURNAL   Patent: EP 0505908-A 12 30-SEP-1992;
            F. HOFFMANN-LA ROCHE AG
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       11 a     24 c      3 g     10 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 6;  Length 48;
  Best Local Similarity   88.9%;  Pred. No. 2.6e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAAGTG 19
          ||||| | ||||||||||
Db     37 GAGGTTCTGGTGGAAGTG 20


RESULT 7
AJ8144
LOCUS       AJ8144                   100 bp    mRNA    linear   PRI 13-NOV-1998
DEFINITION  Homo sapiens NIK gene, exon 6, 5'.
ACCESSION   AJ008144
VERSION     AJ008144.1  GI:3873593
KEYWORDS    NIK gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Brookes,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-1998) A. Brookes, Department of Genetics and
            Pathology, Biomedical Center, Box 589, 751 23 Uppsala, SWEDEN
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Aronsson,F.C., Magnusson,P., Andersson,B., Karsten,S.L.,
            Shibasaki,Y., Lendon,C.L., Goate,A.M. and Brookes,A.J.
  TITLE     The NIK protein kinase and C17orf1 genes: chromosomal mapping, gene
            structures and mutational screening in frontotemporal dementia and
            parkinsonism linked to chromosome 17
  JOURNAL   Hum. Genet. 103 (3), 340-345 (1998)
  MEDLINE   99013435
   PUBMED   9799091
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="q21"
     gene            50. .100
                     /gene="NIK"
     exon            50. .>100
                     /gene="NIK"
                     /number=6
BASE COUNT       16 a     27 c     32 g     25 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 9;  Length 100;
  Best Local Similarity   93.8%;  Pred. No. 3.8e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 AGGTGCGGGTGGAAGT 18
          |||||||| |||||||
Db     49 AGGTGCGGCTGGAAGT 64


RESULT 8
AR106307/c
LOCUS       AR106307                 30 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 11 from patent US 6107025.
ACCESSION   AR106307
VERSION     AR106307.1  GI:12820837
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Caskey,C.Thomas., Nelson,D.L., Pieretti,M., Warren,S.T. and
            Oostra,B.A.
  TITLE     Diagnosis of the fragile X syndrome
  JOURNAL   Patent: US 6107025-A 11 22-AUG-2000;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        6 a     17 c      3 g      4 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAAGTGC 20
          ||| ||  ||||||||||
Db     25 GAGCTGGTGGTGGAAGTGC 7


RESULT 9
AR110116/c
LOCUS       AR110116                 30 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 7 from patent US 6114150.
ACCESSION   AR110116
VERSION     AR110116.1  GI:12826392
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Weissman,S.M. and Baskaran,N.
  TITLE     Amplification of nucleic acids
  JOURNAL   Patent: US 6114150-A 7 05-SEP-2000;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        6 a     17 c      3 g      4 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAAGTGC 20
          ||| ||  ||||||||||
Db     25 GAGCTGGTGGTGGAAGTGC 7


RESULT 10
AR126277/c
LOCUS       AR126277                 30 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 11 from patent US 6180337.
ACCESSION   AR126277
VERSION     AR126277.1  GI:14112870
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Caskey,C.Thomas., Nelson,D.L., Pieretti,M., Warren,S.T.,
            Oostra,B.A. and Fu,Y.-h.
  TITLE     Diagnosis of the fragile X syndrome
```

```
  JOURNAL   Patent: US 6180337-A 11 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        6 a     17 c      3 g      4 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAAGTGC 20
          ||| ||  ||||||||||
Db     25 GAGCTGGTGGTGGAAGTGC 7


RESULT 11
AX161940
LOCUS       AX161940                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5268 from Patent WO0140521.
ACCESSION   AX161940
VERSION     AX161940.1  GI:14543271
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5268 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (5267 is other entry)
                     Accession number cg43988747"
BASE COUNT       11 a     12 c     16 g     12 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 4.8e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAAGTGC 20
          ||| ||  ||||| ||||||
Db     25 GAGTTGTGGGTGTAAGTGC 43


RESULT 12
AF318496/c
LOCUS       AF318496                 85 bp    DNA     linear   INV 22-MAY-2001
DEFINITION  Scutigerella immaculata sex combs reduced protein (Scr) gene,
            partial cds.
ACCESSION   AF318496
VERSION     AF318496.1  GI:12007393
KEYWORDS    .
SOURCE      Scutigerella immaculata.
  ORGANISM  Scutigerella immaculata
            Eukaryota; Metazoa; Arthropoda; Myriapoda; Symphyla;
            Scutigerellidae; Scutigerella.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Cook,C.E., Smith,M.L., Telford,M.J., Bastianello,A. and Akam,M.
  TITLE     Hox genes and the phylogeny of the arthropods
  JOURNAL   Curr. Biol. 11 (10), 759-763 (2001)
  MEDLINE   21272202
   PUBMED   11378385
REFERENCE   2  (bases 1 to 85)
  AUTHORS   Cook,C.E. and Akam,M.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-2000) Museum of Zoology, University of Cambridge,
            Downing Street, Cambridge CB2 3EJ, UK
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Scutigerella immaculata"
                     /db_xref="taxon:146864"
     gene            <1. .>85
                     /gene="Scr"
     CDS             <1. .>85
                     /gene="Scr"
                     /note="homeobox protein"
                     /codon_start=2
                     /product="sex combs reduced protein"
                     /protein_id="AAG45176.1"
                     /db_xref="GI:12007394"
                     /translation="FHFNRYLTRRRRIEIAHSLCLTERQIKI"
BASE COUNT       21 a     30 c     18 g     16 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 3;  Length 85;
  Best Local Similarity   84.2%;  Pred. No. 4.7e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGAGGTGCGGGTGGAAGTG 19
          |||||| | ||| ||||||
Db     23 TGAGGTACCGGTTGAAGTG 5


RESULT 13
A57510/c
LOCUS       A57510                   24 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 2 from Patent WO9632483.
ACCESSION   A57510
VERSION     A57510.1  GI:3713368
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Masucci,M.G.
  TITLE     IMMUNE-EVADING PROTEINS
  JOURNAL   Patent: WO 9632483-A 2 17-OCT-1996;
            MASUCCI MARIA GRAZIA (SE)
COMMENT     Other publication AU 5284296 961030.
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        5 a     14 c      3 g      2 t
ORIGIN

  Query Match             70.0%;  Score 14;  DB 6;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+04;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGA 15
          ||||||||||||||
Db     15 GAGGTGCGGGTGGA 2


RESULT 14
A57516/c
LOCUS       A57516                   24 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 8 from Patent WO9632483.
ACCESSION   A57516
VERSION     A57516.1  GI:3713374
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Masucci,M.G.
  TITLE     IMMUNE-EVADING PROTEINS
```

```
  JOURNAL   Patent: WO 9632483-A 8 17-OCT-1996;
            MASUCCI MARIA GRAZIA (SE)
COMMENT     Other publication AU 5284296 961030.
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        3 a      14 c       2 g       5 t
ORIGIN

  Query Match             70.0%;  Score 14;  DB 6;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+04;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGA 15
           ||||||||||||||
Db      15 GAGGTGCGGGTGGA 2


RESULT 15
A57518/c
LOCUS       A57518                   24 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 10 from Patent WO9632483.
ACCESSION   A57518
VERSION     A57518.1  GI:3713376
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Masucci,M.G.
  TITLE     IMMUNE-EVADING PROTEINS
  JOURNAL   Patent: WO 9632483-A 10 17-OCT-1996;
            MASUCCI MARIA GRAZIA (SE)
COMMENT     Other publication AU 5284296 961030.
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        5 a      13 c       2 g       4 t
ORIGIN

  Query Match             70.0%;  Score 14;  DB 6;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+04;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGA 15
           ||||||||||||||
Db      15 GAGGTGCGGGTGGA 2


Search completed: December  3, 2002, 18:12:56
Job time : 309.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-9
Perfect score: 20
Sequence: 1 tgaggtgcgggtggaagtgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	21	AAZ31439	Human neuropilin m
2		15.2	76.0	60	24	ABN49870	Human spliced tran
c 3		15	75.0	24	19	AAV55815	Multimerisation of
c 4		15	75.0	24	19	AAV55817	Multimerisation of
5		14.8	74.0	48	13	AAQ28729	Sequence of oligon
c 6		14.8	74.0	48	13	AAQ28730	Sequence of oligon
7		14.4	72.0	60	24	ABN48299	Human spliced tran
c 8		14.2	71.0	21	15	AAQ70342	Antisense oligonuc
c 9		14.2	71.0	30	13	AAQ31897	PCR primer #2 for
c	10	14.2	71.0	30	21	AAA95218	G+C rich nucleic a
	11	14.2	71.0	50	20	AAX52367	Probe used to iso
c	12	14.2	71.0	50	22	AAL33769	Human SNP oligonuc
c	13	14.2	71.0	50	22	AAL33771	Human SNP oligonuc
c	14	14.2	71.0	50	22	AAL33772	Human SNP oligonuc
	15	14.2	71.0	50	22	AAF72525	Human PRO polypept
c	16	14.2	71.0	51	22	AAL31016	Human SNP oligonuc
c	17	14.2	71.0	51	22	AAL33770	Human SNP oligonuc
	18	14.2	71.0	51	22	AAI78327	Human silent SNP c
c	19	14.2	71.0	68	24	ABL49587	scFv antibody and
	20	14.2	71.0	77	24	ABL49586	scFv antibody and
	21	14.2	71.0	91	22	AAH47683	Nucleotide sequenc
	22	14.2	71.0	91	24	ABL49551	scFv antibody and
c	23	14	70.0	24	19	AAV55813	Multimerisation of
c	24	14	70.0	24	19	AAV55819	Multimerisation of
c	25	14	70.0	24	19	AAV55821	Multimerisation of
c	26	13.8	69.0	22	20	AAX36787	PCR primer for hum
c	27	13.8	69.0	24	24	AAD29591	Human GABABR3 cDNA
c	28	13.8	69.0	27	21	AAA29250	Antisense primer P
c	29	13.8	69.0	35	21	AAA12990	Cellulomonas fimi
	30	13.8	69.0	38	22	AAS15527	Beta actin gene pr
c	31	13.8	69.0	47	21	AAZ68150	Human map-related
c	32	13.8	69.0	51	22	AAL31328	Human SNP oligonuc
	33	13.8	69.0	60	24	ABN32585	Human spliced tran
	34	13.8	69.0	60	24	ABN40974	Human spliced tran
	35	13.8	69.0	65	24	ABN31223	Rat spliced transc
	36	13.6	68.0	23	24	ABK47083	Mouse EST AA726662
c	37	13.6	68.0	65	21	AAZ92399	Human galectin-1 o
	38	13.6	68.0	65	21	AAZ92400	Human galectin-1 o
	39	13.6	68.0	65	24	ABN31383	Rat spliced transc
c	40	13.6	68.0	96	15	AAQ76782	Human genome fragm
c	41	13.4	67.0	27	19	AAV34033	P. carinii serine
	42	13.4	67.0	97	17	AAT88550	Secretory phosphol
c	43	13.2	66.0	18	15	AAQ70333	Antisense oligonuc
c	44	13.2	66.0	25	22	AAF84027	EPO gene fragment
c	45	13.2	66.0	30	20	AAX27663	DNA encoding a HRG

ALIGNMENTS

RESULT 1
AAZ31439
ID AAZ31439 standard; DNA; 20 BP.
XX
AC AAZ31439;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3609.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 16; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 3 A; 2 C; 11 G; 4 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGGTGCGGGTGGAAGTGC 20
||||||||||||||||||||
Db 1 TGAGGTGCGGGTGGAAGTGC 20

RESULT 2
ABN49870
ID ABN49870 standard; DNA; 60 BP.
XX
AC ABN49870;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:22618.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 22618; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 17 A; 12 C; 18 G; 13 T; 0 other;

Query Match 76.0%; Score 15.2; DB 24; Length 60;
Best Local Similarity 85.0%; Pred. No. 1.2e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGAGGTGCGGGTGGAAGTGC 20
|||||||| |||||| || |
Db 25 TGAGGTGCTGGTGGATGTCC 44

RESULT 3
AAV55815/c
ID AAV55815 standard; DNA; 24 BP.
XX
AC AAV55815;
XX
DT 18-NOV-1998 (first entry)
XX
DE Multimerisation of minimal motifs using primer ZGS2.
XX
KW Fusion protein; stabilising polypeptide; proteolytic degradation;
KW resistance; half-life; autoimmune disease; inflammation; nitro drug;
KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
KW cancer; pathological condition; minimal motif; PCR primer; ss.
XX
OS Synthetic.
OS Epstein-barr virus.
XX
PN WO9822577-A1.
XX
PD 28-MAY-1998.
XX
PF 17-NOV-1997; 97WO-IB01508.
XX
PR 25-JUN-1997; 97US-0048945.
PR 15-NOV-1996; 96US-0030986.
XX
PA (MASU/) MASUCCI M G.
XX
PI Masucci MG;
XX
DR WPI; 1998-312463/27.
XX
PT New fusion proteins resistant to proteolytic degradation -
PT comprising a core protein with a stabilising polypeptide comprising
PT a peptide sequence containing glycine repeats
XX
PS Disclosure; Page 72; 120pp; English.
XX
CC Sequences shown in AAV55812 to AAV55827 represent primers used in the
CC course of the invention for the multimerisation of minimal motifs. The
CC invention provides a method for increasing the resistance of a core
CC protein to proteolytic degradation that comprises linking or inserting
CC onto or into the core protein a stabilising polypeptide of formula

CC [(Glya)X(Glyb)Y(Glyc)Z]n where Glya, Glyb, Glyc are 1-6 sequential Gly
CC residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met, Phe, Pro or Thr
CC and n can be anything between 1-66. X, Y and Z need not be identical
CC from n repeat to n repeat. Alternatively a nucleic acid encoding a
CC stabilising polypeptide can be linked onto or inserted into a nucleic
CC acid encoding a core protein. The fusion proteins of the invention are
CC more resistant to degradation by proteases and, thus, have a longer
CC half-life than the unfused core protein. The products can be used for
CC treating autoimmune diseases, cancer and inflammation. In particular, the
CC core protein may be an IkappaB regulator protein for the treatment of
CC inflammatory bowel disease, or a nitroreductase protein which can
CC activate nitro drugs in enzyme/prodrug therapy to treat cancer or other
CC pathological conditions. The fusion proteins can also be used in
CC diagnostic methods such as in vivo imaging.
XX
SQ Sequence 24 BP; 4 A; 14 C; 2 G; 4 T; 0 other;

Query Match 75.0%; Score 15; DB 19; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GAGGTGCGGGTGGAA 16
||||||||||||||||
Db 15 GAGGTGCGGGTGGAA 1

RESULT 4
AAV55817/c
ID AAV55817 standard; DNA; 24 BP.
XX
AC AAV55817;
XX
DT 18-NOV-1998 (first entry)
XX
DE Multimerisation of minimal motifs using primer ZGR2.
XX
KW Fusion protein; stabilising polypeptide; proteolytic degradation;
KW resistance; half-life; autoimmune disease; inflammation; nitro drug;
KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
KW cancer; pathological condition; minimal motif; PCR primer; ss.
XX
OS Synthetic.
OS Epstein-barr virus.
XX
PN WO9822577-A1.
XX
PD 28-MAY-1998.
XX
PF 17-NOV-1997; 97WO-IB01508.
XX
PR 25-JUN-1997; 97US-0048945.
PR 15-NOV-1996; 96US-0030986.
XX
PA (MASU/) MASUCCI M G.
XX
PI Masucci MG;
XX
DR WPI; 1998-312463/27.
XX
PT New fusion proteins resistant to proteolytic degradation -
PT comprising a core protein with a stabilising polypeptide comprising
PT a peptide sequence containing glycine repeats
XX
PS Disclosure; Page 72; 120pp; English.
XX
CC Sequences shown in AAV55812 to AAV55827 represent primers used in the
CC course of the invention for the multimerisation of minimal motifs. The
CC invention provides a method for increasing the resistance of a core
CC protein to proteolytic degradation that comprises linking or inserting
CC onto or into the core protein a stabilising polypeptide of formula
CC [(Glya)X(Glyb)Y(Glyc)Z]n where Glya, Glyb, Glyc are 1-6 sequential Gly
CC residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met, Phe, Pro or Thr
CC and n can be anything between 1-66. X, Y and Z need not be identical
CC from n repeat to n repeat. Alternatively a nucleic acid encoding a
CC stabilising polypeptide can be linked onto or inserted into a nucleic
CC acid encoding a core protein. The fusion proteins of the invention are
CC more resistant to degradation by proteases and, thus, have a longer
CC half-life than the unfused core protein. The products can be used for
CC treating autoimmune diseases, cancer and inflammation. In particular, the
CC core protein may be an IkappaB regulator protein for the treatment of
CC inflammatory bowel disease, or a nitroreductase protein which can
CC activate nitro drugs in enzyme/prodrug therapy to treat cancer or other
CC pathological conditions. The fusion proteins can also be used in
CC diagnostic methods such as in vivo imaging.
XX
SQ Sequence 24 BP; 3 A; 14 C; 2 G; 5 T; 0 other;

Query Match 75.0%; Score 15; DB 19; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GAGGTGCGGGTGGAA 16
||||||||||||||||
Db 15 GAGGTGCGGGTGGAA 1

RESULT 5
AAQ28729
ID AAQ28729 standard; DNA; 48 BP.
XX
AC AAQ28729;
XX
DT 27-FEB-1993 (first entry)
XX
DE Sequence of oligonucleotide 2119 corresponding to the linking
DE peptide between VH and VL domains of plasmid pFvCD3-H-gamma-3.
XX
KW Chimaeric polypeptide; cell surface CD4 molecule; HIV type 1; HIV-1;
KW human immunodeficiency virus; antigen binding site; cytotoxic T-;
KW cell; CD3 molecule; T-cell receptor complex; PCR; ss.
XX
OS Synthetic.
XX
PN EP505908-A.
XX
PD 30-SEP-1992.
XX
PF 18-MAR-1992; 92EP-0104656.
XX
PR 27-MAR-1991; 91EP-0810220.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Karjalainen K, Lanzavecchia A, Traunecker A;
XX
DR WPI; 1992-325228/40.
XX
PT DNA encoding chimeric bi:specific polypeptide(s) e.g. CD4 and CD3
PT binding site - for treating HIV infection, and also cancers e.g.
PT ovarian carcinoma(s) or leukaemia(s)
XX
PS Example; Page 10; 19pp; English.
XX
CC Plasmid pFvCD3-H-gamma-3 comprises a DNA sequence coding for the
CC variable regions of the heavy and the light chain (VH and VL) of
CC an anti-human CD3 antibody (FvCD3) derived from hybridoma TR66
CC and for all domains except the first of the constant regions of
CC the heavy chain of the human immunoglobulin G3. The second strand
CC synthesis of the sequences coding for the VL and VH-domains of the
CC anti-CD3 hybridoma TR66 was performed by T4 DNA polymerase and 5'
CC oligonucleotides 1570 for VH and 1569 for VL sequences. Final
CC amplification was then achieved by using the above mentioned
CC oligonucleotides and 3' oligonucleotides 1354 and 1239j
CC complementary to mouse gamma-HC (immunoglobulin - gamma - heavy
CC chain) and C-kappa (constant domain of immunoglobulin kappa-light

CC chain) genes. The L region (initiation codon and leader sequence
CC for eukaryotic secretion) of the Sp6 HC gene contg. the small
CC intron was amplified by using 5' olionucleotides 1609 and 3'
CC oligonucleotides 1610 which are also complementary (20 bp) to
CC oligonucleotides 1570 at its 5' end (ie at the beginning of the
CC amplified VH fragment. Junctional and flanking olibonucleotides
CC were designed and used to link the L region from Sp6 VH and VL
CC fragments together in one 35 PCR reaction: the flanking 5' and 3'
CC oligonucleotides were 1609 and 2768 respectively, the latter having
CC complementarity to the end for the mouse J-kappa-5 sequence. The
CC oligonucleotide 2139 which links VH to VL fragments had 20 bp
CC complementarities to the end of the JH2-sequence and to
CC oligonucleotide 1569, ie to the 5' end of the VL sequence. Finally,
CC a synthetic piece of DNA with SalI compatible ends
CC (oligonucleotides 2119 and 2120) corresponding to the linking
CC peptide between VH and VL domains, was cloned into a SalI site of
CC the intermediate plasmid to give plasmid pFvCD3-H-gamma-3.
XX
SQ Sequence 48 BP; 10 A; 3 C; 24 G; 11 T; 0 other;

Query Match 74.0%; Score 14.8; DB 13; Length 48;
Best Local Similarity 88.9%; Pred. No. 1.8e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GAGGTGCGGGTGGAAGTG 19
||||| | ||||||||||
Db 16 GAGGTTCTGGTGGAAGTG 33

RESULT 6
AAQ28730/c
ID AAQ28730 standard; DNA; 48 BP.
XX
AC AAQ28730;
XX
DT 27-FEB-1993 (first entry)
XX
DE Sequence of oligonucleotide 2120 corresponding to the linking
DE peptide between VH and VL domains of plasmid pFvCD3-H-gamma-3.
XX
KW Chimaeric polypeptide; cell surface CD4 molecule; HIV type 1; HIV-1;
KW human immunodeficiency virus; antigen binding site; cytotoxic T-;
KW cell; CD3 molecule; T-cell receptor complex; PCR; ss.
XX
OS Synthetic.
XX
PN EP505908-A.
XX
PD 30-SEP-1992.
XX
PF 18-MAR-1992; 92EP-0104656.
XX
PR 27-MAR-1991; 91EP-0810220.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Karjalainen K, Lanzavecchia A, Traunecker A;
XX
DR WPI; 1992-325228/40.
XX
PT DNA encoding chimeric bi:specific polypeptide(s) e.g. CD4 and CD3
PT binding site - for treating HIV infection, and also cancers e.g.
PT ovarian carcinoma(s) or leukaemia(s)
XX
PS Example; Page 10; 19pp; English.
XX
CC Plasmid pFvCD3-H-gamma-3 comprises a DNA sequence coding for the
CC variable regions of the heavy and the light chain (VH and VL) of
CC an anti-human CD3 antibody (FvCD3) derived from hybridoma TR66
CC and for all domains except the first of the constant regions of
CC the heavy chain of the human immunoglobulin G3. The second strand
CC synthesis of the sequences coding for the VL and VH-domains of the
CC anti-CD3 hybridoma TR66 was performed by T4 DNA polymerase and 5'
CC oligonucleotides 1570 for VH and 1569 for VL sequences. Final
CC amplification was then achieved by using the above mentioned
CC oligonucleotides and 3' oligonucleotides 1354 and 1239j
CC complementary to mouse gamma-HC (immunoglobulin - gamma - heavy
CC chain) and C-kappa (constant domain of immunoglobulin kappa-light
CC chain) genes. The L region (initiation codon and leader sequence
CC for eukaryotic secretion) of the Sp6 HC gene contg. the small
CC intron was amplified by using 5' olionucleotides 1609 and 3'
CC oligonucleotides 1610 which are also complementary (20 bp) to
CC oligonucleotides 1570 at its 5' end (ie at the beginning of the
CC amplified VH fragment. Junctional and flanking olibonucleotides
CC were designed and used to link the L region from Sp6 VH and VL
CC fragments together in one 35 PCR reaction: the flanking 5' and 3'
CC oligonucleotides were 1609 and 2768 respectively, the latter having
CC complementarity to the end for the mouse J-kappa-5 sequence. The
CC oligonucleotide 2139 which links VH to VL fragments had 20 bp
CC complementarities to the end of the JH2-sequence and to
CC oligonucleotide 1569, ie to the 5' end of the VL sequence. Finally,
CC a synthetic piece of DNA with SalI compatible ends
CC (oligonucleotides 2119 and 2120) corresponding to the linking
CC peptide between VH and VL domains, was cloned into a SalI site of
CC the intermediate plasmid to give plasmid pFvCD3-H-gamma-3.
XX
SQ Sequence 48 BP; 11 A; 24 C; 3 G; 10 T; 0 other;

Query Match 74.0%; Score 14.8; DB 13; Length 48;
Best Local Similarity 88.9%; Pred. No. 1.8e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GAGGTGCGGGTGGAAGTG 19
||||| | ||||||||||
Db 37 GAGGTTCTGGTGGAAGTG 20

RESULT 7
ABN48299
ID ABN48299 standard; DNA; 60 BP.
XX
AC ABN48299;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:21047.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 21047; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the

```
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 16 A; 9 C; 24 G; 11 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 24;  Length 60;
  Best Local Similarity   93.8%;  Pred. No. 2.7e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 AGGTGCGGGTGGAAGT 18
           ||||||||||  |||||
Db       6 AGGTGCGGGTTGAAGT 21


RESULT 8
AAQ70342/c
ID   AAQ70342 standard; DNA; 21 BP.
XX
AC   AAQ70342;
XX
DT   15-FEB-1995  (first entry)
XX
DE   Antisense oligonucleotide for mouse FGF.
XX
KW   Fibroblast growth factor; hybridisation; laser procedures;
KW   vascular smooth muscle cell; proliferation;
KW   SMC; vascular stenosis; post angioplasty restenosis;
KW   atherosclerosis; cardiac hypertrophy; organ transplant; ss.
XX
OS   Synthetic.
XX
PN   WO9415945-A.
XX
PD   21-JUL-1994.
XX
PF   28-DEC-1993;   93WO-US12600.
XX
PR   31-DEC-1992;   92US-0999706.
XX
PA   (TEXA-) TEXAS BIOTECHNOLOGY CORP.
XX
PI   Denner LA,  Dixon RAF,  Rege AA,  Dixon RA;
XX
DR   WPI; 1994-249123/30.
XX
PT   New anti-sense polynucleotide(s) to fibroblast growth factor
PT   receptor - used for inhibiting vascular smooth muscle cell
PT   proliferation, partic. for treating restenosis
XX
PS   Claim 3; Page 9; 53pp; English.
XX
CC   The sequence is an antisense molecule directed against position -3
CC   to +18, relative to the start codon of the gene for
CC   mouse fibroblast growth factor 1. The polynucleotide can be used for
CC   inhibiting vascular smooth muscle cell proliferation and for treating
CC   a disease e.g. vascular stenosis, post angioplasty restenosis,
CC   atherectomy, atherosclerosis, atrial venous shunt failure, cardiac
CC   hypertrophy, vascular surgery and organ transplant.
CC   See also AAQ70333-60.
XX
SQ   Sequence 21 BP; 4 A; 12 C; 2 G; 3 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 15;  Length 21;
  Best Local Similarity   84.2%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTGC 20
           || ||| || |||||||||
Db      19 GATGTGGGGCTGGAAGTGC 1


RESULT 9
AAQ31897/c
ID   AAQ31897 standard; cDNA; 30 BP.
XX
AC   AAQ31897;
XX
DT   22-APR-1993  (first entry)
XX
DE   PCR primer #2 for X-linked disease detection by (CGG)n repeat detn.
XX
KW   Fragile X disease; sex chromosome; X chromosome; X linked syndrome;
KW   X linked retardation; X linked manic depression; TKCR;
KW   Martin-Bell syndrome; CA polymorphisms; PCR analysis; ss.
XX
OS   Homo sapiens.
XX
PN   WO9220825-A.
XX
PD   26-NOV-1992.
XX
PF   22-MAY-1992;   92WO-US04447.
XX
PR   24-MAY-1991;   91US-0705490.
PR   29-AUG-1991;   91US-0751891.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
PA   (UYEM-) UNIV EMORY SCHOOL MEDICINE.
XX
PI   Caskey CT,  Nelson DL,  Oostra BA,  Pieretti M,  Warren ST;
XX
DR   WPI; 1992-415801/50.
XX
PT   Gene sequence, related probes and cosmid(s) - useful in
PT   diagnosing fragile X syndrome, X-linked metal retardation, manic
PT   depression, and Martin Bell syndrome
XX
PS   Claim 25; Page 54; 75pp; English.
XX
CC   This PCR primer is used together with Q318975 to determine the
CC   variation of the (CGG)n repeat found at the 5'end of the FMR-1
CC   gene. n for normal is in the range 16-30 and for X linked diseases,
CC   n is greater than 30. In the case of fragile X, n is usually twice
CC   the normal range. Othe disease that can be detected include
CC   X-linked mental retardation (of both fragile X and non-fragile X
CC   type), X linked manic depressive disease, TKCR syndrome and
CC   Martin-Bell syndrome.
XX
SQ   Sequence 30 BP; 6 A; 17 C; 3 G; 4 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 13;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTGC 20
           ||| ||  ||||||||||
```

```
Db     25 GAGCTGGTGGTGGAAGTGC 7


RESULT 10
AAA95218/c
ID   AAA95218 standard; DNA; 30 BP.
XX
AC   AAA95218;
XX
DT   12-JAN-2001  (first entry)
XX
DE   G+C rich nucleic acid PCR primer f.
XX
KW   GC rich nucleic acid; DNA amplification; zwitterion;
KW   base pairing disruption; polymerase chain reaction; fragile X syndrome;
KW   spinobulbar muscular atrophy; myotonic dystrophy; Huntington's disease;
KW   Huntington's chorea; spinocerebellar ataxia type 1; PCR primer; ss.
XX
OS   Synthetic.
XX
PN   US6114150-A.
XX
PD   05-SEP-2000.
XX
PF   02-DEC-1996;   96US-0758662.
XX
PR   29-NOV-1995;   95US-0564653.
XX
PA   (UYYA ) UNIV YALE.
XX
PI   Weissman SM,  Baskaran N;
XX
DR   WPI; 2000-586478/55.
XX
PT   Uniform amplification of heterogeneous mixture of nucleic acid
PT   templates of varying G+C content, comprises amplifying the reaction
PT   mixture in the presence of betaine and dimethyl sulphoxide  -
XX
PS   Example 5; column 10; 15pp; English.
XX
CC   The present sequence is a PCR primer used to demonstrate the method of
CC   the invention, which involves a novel way of amplifying GC rich nucleic
CC   acids. In addition to the normal mixture a zwitterion and a compound
CC   which disrupts base pairing are added to the PCR reaction. The zwitterion
CC   is preferably betaine (trimethyl glycine), D-carnitine, dimethyl glycine
CC   or monomethyl glycine, and the base pair disruption compound may be DMSO,
CC   formamide, sodium perchlorate or glyoxyl, among others. This method is
CC   useful as it allows the uniform and more efficient amplification of
CC   nucleic acids composed of GC rich regions. This is useful in DNA library
CC   construction, as the 5' end of genes are often GC rich, and in the
CC   diagnosis of diseases where GC rich triplets are expanded, for example
CC   fragile X syndrome, spinobulbar muscular atrophy, myotonic dystrophy,
CC   Huntington's disease and spinocerebellar ataxia type 1.
XX
SQ   Sequence 30 BP; 6 A; 17 C; 3 G; 4 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 21;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAAGTGC 20
          ||| ||  ||||||||||
Db     25 GAGCTGGTGGTGGAAGTGC 7


RESULT 11
AAX52367
ID   AAX52367 standard; DNA; 50 BP.
XX
AC   AAX52367;
XX
DT   25-JUN-1999  (first entry)
XX
DE   Probe  used to isolate a cDNA clone encoding PRO229.
XX
KW   Secreted protein; transmembrane protein; human; enterocolitis;
KW   Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW   congenital microvillus atrophy; skin disease; cell growth;
KW   abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW   Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW   fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
KW   anti-thrombotic; wound healing; tissue repair; probe; ss.
XX
OS   Synthetic.
XX
PN   WO9914328-A2.
XX
PD   25-MAR-1999.
XX
PF   16-SEP-1998;   98WO-US19330.
XX
PR   25-NOV-1997;   97US-0066840.
PR   17-SEP-1997;   97US-0059113.
PR   17-SEP-1997;   97US-0059115.
PR   17-SEP-1997;   97US-0059117.
PR   17-SEP-1997;   97US-0059119.
PR   17-SEP-1997;   97US-0059121.
PR   17-SEP-1997;   97US-0059122.
PR   17-SEP-1997;   97US-0059184.
PR   18-SEP-1997;   97US-0059263.
PR   18-SEP-1997;   97US-0059266.
PR   15-OCT-1997;   97US-0062125.
PR   17-OCT-1997;   97US-0062285.
PR   17-OCT-1997;   97US-0062287.
PR   21-OCT-1997;   97US-0063486.
PR   24-OCT-1997;   97US-0062814.
PR   24-OCT-1997;   97US-0062816.
PR   24-OCT-1997;   97US-0063045.
PR   24-OCT-1997;   97US-0063120.
PR   24-OCT-1997;   97US-0063121.
PR   24-OCT-1997;   97US-0063127.
PR   24-OCT-1997;   97US-0063128.
PR   27-OCT-1997;   97US-0063329.
PR   27-OCT-1997;   97US-0063327.
PR   28-OCT-1997;   97US-0063541.
PR   28-OCT-1997;   97US-0063542.
PR   28-OCT-1997;   97US-0063544.
PR   28-OCT-1997;   97US-0063549.
PR   28-OCT-1997;   97US-0063550.
PR   28-OCT-1997;   97US-0063564.
PR   29-OCT-1997;   97US-0063435.
PR   29-OCT-1997;   97US-0063704.
PR   29-OCT-1997;   97US-0063732.
PR   29-OCT-1997;   97US-0063738.
PR   29-OCT-1997;   97US-0063734.
PR   29-OCT-1997;   97US-0064215.
PR   29-OCT-1997;   97US-0063735.
PR   31-OCT-1997;   97US-0063870.
PR   31-OCT-1997;   97US-0064103.
PR   03-NOV-1997;   97US-0064248.
PR   07-NOV-1997;   97US-0064809.
PR   12-NOV-1997;   97US-0065186.
PR   17-NOV-1997;   97US-0065846.
PR   18-NOV-1997;   97US-0065693.
PR   21-NOV-1997;   97US-0066120.
PR   21-NOV-1997;   97US-0066364.
PR   24-NOV-1997;   97US-0066772.
PR   24-NOV-1997;   97US-0066466.
PR   24-NOV-1997;   97US-0066770.
PR   24-NOV-1997;   97US-0066511.
PR   24-NOV-1997;   97US-0066453.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Chen J,  Goddard A,  Gurney AL,  Pennica D,  Wood WI,  Yuan J;
XX
```

DR WPI; 1999-229533/19.
XX
PT New isolated human genes and polypeptides used in, e.g. treatment of
PT gastrointestinal ulceration
XX
PS Example 24; Page 133; 320pp; English.
XX
CC Oligonucleotides AAX52276-532 represent PCR primers and probes used
CC to isolate and amplify cDNA encoding secreted and transmembrane
CC human proteins (see AAX52213-74 and AAY13344-403). The cDNA sequences
CC are obtained from cDNA libraries, prepared from fetal lung,
CC fetal kidney, fetal brain, fetal liver and fetal retina. The encoded
CC polypeptides have specific uses based on their homology to known
CC polypeptides, e.g. PRO211 and PRO217 can be used for disorders
CC associated with the preservation and maintenance of gastrointestinal
CC mucosa and the repair of acute and chronic mucosal lesions (e.g.
CC enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration
CC and congenital microvillus atrophy), skin diseases associated with
CC abnormal keratinocyte differentiation (e.g. psoriasis, epithelial cancers
CC such as lung squamous cell carcinoma of the vulva and gliomas), potent
CC effects on cell growth and development, diseases related to growth or
CC survival of nerve cells including Parkinson's disease, Alzheimer's
CC disease, ALS, neuropathies or cancer. PRO265 can be used as for
CC fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used as a
CC target for anti-tumor drugs. PRO533 may be used in the treatment of Usher
CC Syndrome or Atrophia areata; PRO269 can be used as an anti-thrombotic
CC agent; PRO287 polypeptides and portions may have therapeutic applications
CC in wound healing and tissue repair; PRO317 can be used for treating
CC problems of the kidney, uterus, endometrium, blood vessels, or related
CC tissue, e.g. in the heart of genital tract.
XX
SQ Sequence 50 BP; 11 A; 11 C; 22 G; 6 T; 0 other;

Query Match 71.0%; Score 14.2; DB 20; Length 50;
Best Local Similarity 84.2%; Pred. No. 3.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TGAGGTGCGGGTGGAAGTG 19
           ||| | ||||||||| |||
Db      16 TGAAGGGCGGGTGGAGGTG 34
```

RESULT 12
AAL33769/c
ID AAL33769 standard; DNA; 50 BP.
XX
AC AAL33769;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #6977.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 3381; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 50 BP; 8 A; 25 C; 8 G; 9 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 50;
Best Local Similarity 84.2%; Pred. No. 3.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 GAGGTGCGGGTGGAAGTGC 20
           | |||| |||||| |||||
Db      49 GGGGTGAGGGTGGTAGTGC 31
```

RESULT 13
AAL33771/c
ID AAL33771 standard; DNA; 50 BP.
XX
AC AAL33771;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #6979.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX

DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 3382; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 50 BP; 7 A; 26 C; 7 G; 10 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 50;
Best Local Similarity 84.2%; Pred. No. 3.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GAGGTGCGGGTGGAAGTGC 20
| |||| |||||| |||||
Db 45 GGGGTGAGGGTGGTAGTGC 27

RESULT 14
AAL33772/c
ID AAL33772 standard; DNA; 50 BP.
XX
AC AAL33772;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #6980.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 3382; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 50 BP; 8 A; 26 C; 7 G; 9 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 50;
Best Local Similarity 84.2%; Pred. No. 3.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GAGGTGCGGGTGGAAGTGC 20
| |||| |||||| |||||
Db 43 GGGGTGAGGGTGGTAGTGC 25

RESULT 15
AAF72525
ID AAF72525 standard; DNA; 50 BP.
XX
AC AAF72525;
XX
DT 24-APR-2001 (first entry)
XX
DE Human PRO polypeptide gene hybridisation probe SEQ ID NO: 151.
XX
KW Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
KW antiparkinsonian nootropic; neuroprotective; vulnerary; cardiant;
KW antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
KW antiarthritic; antiinfertility; antidiabetic; antiviral; diabetes;
KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW ischaemia; inflammation; probe; ss.
XX
OS Homo sapiens.
XX
PN WO200104311-A1.
XX
PD 18-JAN-2001.
XX
PF 22-FEB-2000; 2000WO-US04414.
XX
PR 07-JUL-1999; 99US-0143048.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 99WO-US00219.
XX

PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2001-081051/09.
XX
PT Sixty one nucleic acids encoding PRO polypeptides which are useful in
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
PT Alzheimer's disease) -
XX
PS Example 24; Page 170; 393pp; English.
XX
CC The present sequence is a probe which was used in the isolation of one
CC of sixty one nucleic acids encoding novel secreted and transmembrane PRO
CC polypeptides. The PRO polypeptides are useful for treating skin diseases
CC (e.g. psoriasis), cancers (e.g. lung squamous cell carcinoma),
CC gastrointestinal disorders (e.g. enterocolitis), neurodegenerative
CC diseases (e.g. Alzheimer's disease, Parkinson's disease), wound repair,
CC cardiovascular disorders (e.g. endometrial bleeding angiogenesis,
CC ischaemias such as coronary ischaemia, atherosclerosis), inflammatory
CC disorders (e.g. asthma, rheumatoid arthritis, multiple sclerosis),
CC infertility, AIDS and diabetes and retinal disorders such as retinitis
CC pigmentosum. The PRO nucleic acids have applications in molecular
CC biology, including use as hybridization probes, and in chromosome and
CC gene mapping.
XX
SQ Sequence 50 BP; 11 A; 11 C; 22 G; 6 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 50;
Best Local Similarity 84.2%; Pred. No. 3.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TGAGGTGCGGGTGGAAGTG 19
           ||| | ||||||||| |||
Db      16 TGAAGGGCGGGTGGAGGTG 34
```

Search completed: November 23, 2002, 06:28:55
Job time : 101.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-9
Perfect score: 20
Sequence: 1 tgaggtgcgggtggaagtgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	15	75.0	24	2	US-08-529-190B-7	Sequence 7, Appli
c	2	15	75.0	24	2	US-08-529-190B-10	Sequence 10, Appl
c	3	14.2	71.0	21	5	PCT-US93-12600-10	Sequence 10, Appl
c	4	14.2	71.0	30	3	US-07-705-490-11	Sequence 11, Appl
c	5	14.2	71.0	30	3	US-08-758-662-7	Sequence 7, Appli
c	6	14.2	71.0	30	4	US-07-751-891B-11	Sequence 11, Appl
c	7	14	70.0	24	2	US-08-529-190B-4	Sequence 4, Appli
c	8	14	70.0	24	2	US-08-529-190B-13	Sequence 13, Appl
c	9	14	70.0	24	2	US-08-529-190B-16	Sequence 16, Appl
c	10	13.8	69.0	48	1	US-07-917-034A-20	Sequence 20, Appl
	11	13.4	67.0	97	1	US-08-399-412A-79	Sequence 79, Appl
c	12	13.2	66.0	18	5	PCT-US93-12600-1	Sequence 1, Appli
	13	13.2	66.0	30	4	US-08-506-296B-34	Sequence 34, Appl
	14	13.2	66.0	51	1	US-08-285-936-44	Sequence 44, Appl
c	15	13.2	66.0	51	1	US-08-285-936-45	Sequence 45, Appl
	16	13.2	66.0	51	1	US-08-487-860-44	Sequence 44, Appl
c	17	13.2	66.0	51	1	US-08-487-860-45	Sequence 45, Appl
	18	12.8	64.0	33	4	US-09-206-059-70	Sequence 70, Appl
	19	12.8	64.0	60	5	PCT-US95-11985A-4	Sequence 4, Appli
	20	12.8	64.0	69	2	US-08-894-578-149	Sequence 149, App
c	21	12.6	63.0	20	2	US-08-837-201C-2	Sequence 2, Appli
c	22	12.6	63.0	20	4	US-09-364-416-2	Sequence 2, Appli
c	23	12.6	63.0	21	5	PCT-US93-12600-25	Sequence 25, Appl
	24	12.6	63.0	23	4	US-09-262-773-203	Sequence 203, App
c	25	12.6	63.0	24	5	PCT-US93-12600-23	Sequence 23, Appl
c	26	12.6	63.0	24	5	PCT-US93-12600-24	Sequence 24, Appl
c	27	12.6	63.0	25	4	US-09-009-816-10	Sequence 10, Appl
c	28	12.6	63.0	27	5	PCT-US93-12600-22	Sequence 22, Appl
c	29	12.6	63.0	40	3	US-09-277-016-14	Sequence 14, Appl
	30	12.6	63.0	42	1	US-07-640-029-9	Sequence 9, Appli
	31	12.6	63.0	42	1	US-07-921-807B-15	Sequence 15, Appl
	32	12.6	63.0	42	1	US-08-441-944A-15	Sequence 15, Appl
	33	12.6	63.0	42	4	US-08-439-992A-9	Sequence 9, Appli
c	34	12.6	63.0	45	1	US-08-160-670A-33	Sequence 33, Appl
c	35	12.6	63.0	50	1	US-08-171-389-474	Sequence 474, App
c	36	12.6	63.0	50	1	US-08-123-936-474	Sequence 474, App
c	37	12.6	63.0	50	2	US-08-475-228A-474	Sequence 474, App
c	38	12.6	63.0	50	3	US-08-482-080A-474	Sequence 474, App
c	39	12.6	63.0	50	4	US-09-354-947-474	Sequence 474, App
c	40	12.6	63.0	50	5	PCT-US93-12388-474	Sequence 474, App
c	41	12.6	63.0	61	2	US-08-771-624B-19	Sequence 19, Appl
	42	12.4	62.0	17	4	US-09-420-211-15	Sequence 15, Appl
	43	12.4	62.0	21	3	US-08-891-789B-26	Sequence 26, Appl
c	44	12.4	62.0	36	1	US-08-411-795B-414	Sequence 414, App
	45	12.4	62.0	36	1	US-08-411-795B-415	Sequence 415, App

ALIGNMENTS

RESULT 1
US-08-529-190B-7/c
; Sequence 7, Application US/08529190B
; Patent No. 5833991
; GENERAL INFORMATION:
; APPLICANT: Masucci, Maria G.
; TITLE OF INVENTION: GLYCINE-CONTAINING SEQUENCES
; TITLE OF INVENTION: CONFERRING INVISIBILITY TO THE IMMUNE SYSTEM
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: One Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,190B
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE9501324-9
; FILING DATE: 10-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US08/522,595
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen A
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 3255/53015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
US-08-529-190B-7

Query Match 75.0%; Score 15; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 GAGGTGCGGGTGGAA 16
          ||||||||||||||||
Db     15 GAGGTGCGGGTGGAA 1


RESULT 2
US-08-529-190B-10/c
; Sequence 10, Application US/08529190B
; Patent No. 5833991
;  GENERAL INFORMATION:
;    APPLICANT:  Masucci, Maria G.
;    TITLE OF INVENTION:  GLYCINE-CONTAINING SEQUENCES
;    TITLE OF INVENTION:  CONFERRING INVISIBILITY TO THE IMMUNE SYSTEM
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  One Financial Center
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Wordperfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/529,190B
;      FILING DATE:  15-SEP-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  SE9501324-9
;      FILING DATE:  10-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US08/522,595
;      FILING DATE:  01-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Ph.D., Kathleen A
;      REGISTRATION NUMBER:  34,380
;      REFERENCE/DOCKET NUMBER:  3255/53015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-345-9100
;      TELEFAX:  617-345-9111
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-529-190B-10

  Query Match             75.0%;  Score 15;  DB 2;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAA 16
          ||||||||||||||||
Db     15 GAGGTGCGGGTGGAA 1


RESULT 3
PCT-US93-12600-10/c
; Sequence 10, Application PC/TUS9312600
;  GENERAL INFORMATION:
;    APPLICANT:  Denner, Larry A.
;    APPLICANT:  Rege, Ajay A.
;    APPLICANT:  Dixon, Richard A.F.
;    TITLE OF INVENTION:  ANTISENSE MOLECULES DIRECTED AGAINST A
;    TITLE OF INVENTION:  FIBROBLAST GROWTH FACTOR RECEPTOR GENE FAMILY
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dressler, Goldsmith, Shore &
;      ADDRESSEE:  Milnamow, Ltd.
;      STREET:  180 North Stetson, Suite 4700
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12600
;      FILING DATE:  28-DEC-1993
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/999,706
;      FILING DATE:  December 31, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Katz, Martin L.
;      REGISTRATION NUMBER:  25,011
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312)616-5400
;      TELEFAX:  (312)616-5460
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
PCT-US93-12600-10

  Query Match             71.0%;  Score 14.2;  DB 5;  Length 21;
  Best Local Similarity   84.2%;  Pred. No. 4e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GAGGTGCGGGTGGAAGTGC 20
          || ||| || |||||||||
Db     19 GATGTGGGGCTGGAAGTGC 1


RESULT 4
US-07-705-490-11/c
; Sequence 11, Application US/07705490
; Patent No. 6107025
;  GENERAL INFORMATION:
;    APPLICANT:  Caskey, C. T.
;    APPLICANT:  Nelson, David L.
;    APPLICANT:  Pieretti, Maura
;    APPLICANT:  Warren, Stephen T.
;    APPLICANT:  Oostra, Ben A.
;    TITLE OF INVENTION:  Diagnosis of the Fragile X Syndrome
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Thomas D. Paul
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.A.
;      ZIP:  77010-3095
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/705,490
;      FILING DATE:  19910708
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, Thomas D.
;      REGISTRATION NUMBER:  32,714
```

```
;       REFERENCE/DOCKET NUMBER:  D-5350
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  713/651-5325
;       TELEFAX:  713/651-5246
;       TELEX:  762829
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  30 base pairs
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
US-07-705-490-11

  Query Match             71.0%;  Score 14.2;  DB 3;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 4.1e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTGC 20
           ||| ||  ||||||||||
Db      25 GAGCTGGTGGTGGAAGTGC 7


RESULT 5
US-08-758-662-7/c
; Sequence 7, Application US/08758662
; Patent No. 6114150
;  GENERAL INFORMATION:
;     APPLICANT:  Weissman, Sherman
;     APPLICANT:  Baskaran, Namadev
;     TITLE OF INVENTION:  AMPLIFICATION OF NUCLEIC ACIDS
;     NUMBER OF SEQUENCES:  8
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  SEED and BERRY LLP
;       STREET:  701 Fifth Avenue, 6300 Columbia Center
;       CITY:  Seattle
;       STATE:  Washington
;       COUNTRY:  USA
;       ZIP:  98104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/758,662
;       FILING DATE:  29-NOV-1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 6114150tenburg Ph.D., Carol
;       REGISTRATION NUMBER:  39,317
;       REFERENCE/DOCKET NUMBER:  390036.402C1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  30 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-758-662-7

  Query Match             71.0%;  Score 14.2;  DB 3;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 4.1e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTGC 20
           ||| ||  ||||||||||
Db      25 GAGCTGGTGGTGGAAGTGC 7


RESULT 6
US-07-751-891B-11/c
; Sequence 11, Application US/07751891B
; Patent No. 6180337
;   GENERAL INFORMATION:
;       APPLICANT: Caskey, C. T.
;                  Nelson, David L.
;                  Pieretti, Maura
;                  Warren, Stephen T.
;                  Oostra, Ben A.
;                  Fu, Ying-hui
;       TITLE OF INVENTION: Diagnosis of the Fragile X Syndrome
;       NUMBER OF SEQUENCES: 24
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Thomas D. Paul
;            STREET: 1301 McKinney, Suite 5100
;            CITY: Houston
;            STATE: Texas
;            COUNTRY: U.S.A.
;            ZIP: 77010-3095
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/751,891B
;            FILING DATE: 29-Aug-1991
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Paul, Thomas D.
;            REGISTRATION NUMBER: 32,714
;            REFERENCE/DOCKET NUMBER: D-5350
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 713/651-5325
;            TELEFAX: 713/651-5246
;            TELEX: 762829
;   INFORMATION FOR SEQ ID NO: 11:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 30 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       HYPOTHETICAL: NO
;       SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-07-751-891B-11

  Query Match             71.0%;  Score 14.2;  DB 4;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 4.1e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTGC 20
           ||| ||  ||||||||||
Db      25 GAGCTGGTGGTGGAAGTGC 7


RESULT 7
US-08-529-190B-4/c
; Sequence 4, Application US/08529190B
; Patent No. 5833991
;  GENERAL INFORMATION:
;    APPLICANT:  Masucci, Maria G.
;    TITLE OF INVENTION:  GLYCINE-CONTAINING SEQUENCES
;    TITLE OF INVENTION:  CONFERRING INVISIBILITY TO THE IMMUNE SYSTEM
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  One Financial Center
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02111
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Wordperfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/529,190B
;      FILING DATE:  15-SEP-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  SE9501324-9
;      FILING DATE:  10-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US08/522,595
;      FILING DATE:  01-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Ph.D., Kathleen A
;      REGISTRATION NUMBER:  34,380
;      REFERENCE/DOCKET NUMBER:  3255/53015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-345-9100
;      TELEFAX:  617-345-9111
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
US-08-529-190B-4

  Query Match             70.0%;  Score 14;  DB 2;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+02;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGA 15
           ||||||||||||||
Db      15 GAGGTGCGGGTGGA 2


RESULT 8
US-08-529-190B-13/c
; Sequence 13, Application US/08529190B
; Patent No. 5833991
;  GENERAL INFORMATION:
;    APPLICANT:  Masucci, Maria G.
;    TITLE OF INVENTION:  GLYCINE-CONTAINING SEQUENCES
;    TITLE OF INVENTION:  CONFERRING INVISIBILITY TO THE IMMUNE SYSTEM
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  One Financial Center
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Wordperfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/529,190B
;      FILING DATE:  15-SEP-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  SE9501324-9
;      FILING DATE:  10-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US08/522,595
;      FILING DATE:  01-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Ph.D., Kathleen A
;      REGISTRATION NUMBER:  34,380
;      REFERENCE/DOCKET NUMBER:  3255/53015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-345-9100
;      TELEFAX:  617-345-9111
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
US-08-529-190B-13

  Query Match             70.0%;  Score 14;  DB 2;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+02;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGA 15
           ||||||||||||||
Db      15 GAGGTGCGGGTGGA 2


RESULT 9
US-08-529-190B-16/c
; Sequence 16, Application US/08529190B
; Patent No. 5833991
;  GENERAL INFORMATION:
;    APPLICANT:  Masucci, Maria G.
;    TITLE OF INVENTION:  GLYCINE-CONTAINING SEQUENCES
;    TITLE OF INVENTION:  CONFERRING INVISIBILITY TO THE IMMUNE SYSTEM
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  One Financial Center
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  Wordperfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/529,190B
;      FILING DATE:  15-SEP-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  SE9501324-9
;      FILING DATE:  10-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US08/522,595
;      FILING DATE:  01-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Ph.D., Kathleen A
;      REGISTRATION NUMBER:  34,380
;      REFERENCE/DOCKET NUMBER:  3255/53015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-345-9100
;      TELEFAX:  617-345-9111
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
US-08-529-190B-16

  Query Match             70.0%;  Score 14;  DB 2;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+02;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       2 GAGGTGCGGGTGGA 15
           ||||||||||||||
Db      15 GAGGTGCGGGTGGA 2


RESULT 10
US-07-917-034A-20/c
; Sequence 20, Application US/07917034A
; Patent No. 5427927
;  GENERAL INFORMATION:
;    APPLICANT:  Meyer, Thomas, Pohlner, Johannes, Schumacher,
;    APPLICANT:  G nter, Dony, Carola
;    TITLE OF INVENTION:  Process For The Enzymatic Cleavage
;    TITLE OF INVENTION:  of Recombinant Proteins Using IgA Proteases
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/917,034A
;      FILING DATE:  19920830
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP91/00192
;      FILING DATE:  1-Feb-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 40 39 415.8
;      FILING DATE:  10-Dec-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 40 15 922.1
;      FILING DATE:  17-May-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 40 15 921.3
;      FILING DATE:  17-May-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 40 03 149.7
;      FILING DATE:  3-Feb-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5427927man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  HUBR 1018
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-917-034A-20

  Query Match             69.0%;  Score 13.8;  DB 1;  Length 48;
  Best Local Similarity   88.2%;  Pred. No. 6.3e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 AGGTGCGGGTGGAAGTG 19
           |||||| |||||| |||
Db      30 AGGTGCTGGTGGACGTG 14


RESULT 11
US-08-399-412A-79
; Sequence 79, Application US/08399412A
; Patent No. 5622828
;  GENERAL INFORMATION:
;    APPLICANT:  Parma, David
;    APPLICANT:  Gold, Larry
;    TITLE OF INVENTION:  High-Affinity Oligonucleotide
;    TITLE OF INVENTION:  Ligands To Secretory Phospholipase
;    TITLE OF INVENTION:  A2 (sPLA2)
;    NUMBER OF SEQUENCES:  122
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 MB storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,412A
;      FILING DATE:  6-MARCH-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Julie L. Bernard
;      REGISTRATION NUMBER:  36,450
;      REFERENCE/DOCKET NUMBER:  NEX27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  79:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  97 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:  ALL C'S ARE 2'-NH2 cytosine
;    FEATURE:
;      OTHER INFORMATION:  ALL U'S ARE 2'-NH2 uracil
US-08-399-412A-79

  Query Match             67.0%;  Score 13.4;  DB 1;  Length 97;
  Best Local Similarity   73.3%;  Pred. No. 1e+03;
  Matches   11;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 GTGCGGGTGGAAGTG 19
           |:| |||:|||||:|
Db      42 GUGUGGGUGGAAGUG 56


RESULT 12
PCT-US93-12600-1/c
; Sequence 1, Application PC/TUS9312600
;  GENERAL INFORMATION:
;    APPLICANT:  Denner, Larry A.
;    APPLICANT:  Rege, Ajay A.
;    APPLICANT:  Dixon, Richard A.F.
;    TITLE OF INVENTION:  ANTISENSE MOLECULES DIRECTED AGAINST A
;    TITLE OF INVENTION:  FIBROBLAST GROWTH FACTOR RECEPTOR GENE FAMILY
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dressler, Goldsmith, Shore &
;      ADDRESSEE:  Milnamow, Ltd.
```

```
;       STREET:  180 North Stetson, Suite 4700
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60601
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/12600
;       FILING DATE:  28-DEC-1993
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/999,706
;       FILING DATE:  December 31, 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Katz, Martin L.
;       REGISTRATION NUMBER:  25,011
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312)616-5400
;       TELEFAX:  (312)616-5460
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  18 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
PCT-US93-12600-1

  Query Match             66.0%;  Score 13.2;  DB 5;  Length 18;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 AGGTGCGGGTGGAAGTGC 20
           | ||| || |||||||||
Db      18 ATGTGGGGCTGGAAGTGC 1


RESULT 13
US-08-506-296B-34
; Sequence 34, Application US/08506296B
; Patent No. 6313265
;  GENERAL INFORMATION:
;     APPLICANT:  Phillips, Greg
;     APPLICANT:  Cunningham, Bruce A.
;     APPLICANT:  Crossin, Kathryn L.
;     TITLE OF INVENTION:  NEURITE OUTGROWTH-PROMOTING POLYPEPTIDES
;     TITLE OF INVENTION:  CONTAINING FIBRONECTIN TYPE III REPEATS AND METHODS OF USE
;     NUMBER OF SEQUENCES:  77
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  The Scripps Research Institute
;       STREET:  10550 No. 6313265th Torrey Pines Road, TPC-8
;       CITY:  La Jolla
;       STATE:  California
;       COUNTRY:  U.S.
;       ZIP:  92037
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/506,296B
;       FILING DATE:  24-JUL-1995
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fitting, Thomas
;       REGISTRATION NUMBER:  34,163
;       REFERENCE/DOCKET NUMBER:  TSRI 488.0
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (619) 554-2937
;       TELEFAX:  (619) 554-6312
;   INFORMATION FOR SEQ ID NO:  34:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  30 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-506-296B-34

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 30;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 AGGTGCGGGTGGAAGTGC 20
           | ||| |||||||| |||
Db      11 ATGTGGGGGTGGAACTGC 28


RESULT 14
US-08-285-936-44
; Sequence 44, Application US/08285936
; Patent No. 5728821
;  GENERAL INFORMATION:
;     APPLICANT:  Yelton, Dale
;     APPLICANT:  Glaser, Scott
;     APPLICANT:  Huse, William
;     APPLICANT:  Rosok, Mae J.
;     TITLE OF INVENTION:  No. 5728821el Mutant BR96 Antibodies and
;     TITLE OF INVENTION:  Functional Equivalents Reactive With Human Carcinomas
;     NUMBER OF SEQUENCES:  57
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Merchant & Gould
;       STREET:  11150 Santa Monica Bldv., Suite 400
;       CITY:  Los Angeles
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  90025-3395
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/285,936
;       FILING DATE:  04-AUG-1994
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Adriano, Sarah B.
;       REGISTRATION NUMBER:  34,470
;       REFERENCE/DOCKET NUMBER:  30436.16US01
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  310-445-1140
;       TELEFAX:  310-445-9031
;   INFORMATION FOR SEQ ID NO:  44:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  51 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-285-936-44

  Query Match             66.0%;  Score 13.2;  DB 1;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGGTGCGGGTGGAAGTG 19
           ||||| | |||||| |||
Db      13 GAGGTTCTGGTGGAGGTG 30
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-1
Perfect score: 20
Sequence: 1 gagcggcagccccctctcca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		15.4	77.0	57	6	AX286668	AX286668 Sequence
2	c	14.8	74.0	60	6	A00424	A00424 Nucleotide
3	c	14.8	74.0	80	5	CHKER5LD	M13372 Chicken RAV
4		14.4	72.0	55	6	AX073623	AX073623 Sequence
5		14.4	72.0	55	6	AX359862	AX359862 Sequence
6		14.4	72.0	60	6	AX286671	AX286671 Sequence
7	c	14.2	71.0	51	6	AX204378	AX204378 Sequence
8	c	14	70.0	18	6	AX481058	AX481058 Sequence
9	c	13.6	68.0	55	6	AR194772	AR194772 Sequence
10	c	13.6	68.0	100	12	P531	K02977 Plasmid pMM
11		13.4	67.0	31	6	AX248931	AX248931 Sequence
12	c	13.4	67.0	37	6	AX207259	AX207259 Sequence
13	c	13.4	67.0	42	6	AX207261	AX207261 Sequence
14		13.4	67.0	50	6	AX161712	AX161712 Sequence
15	c	13.4	67.0	51	6	AX157115	AX157115 Sequence
16	c	13.4	67.0	51	6	AX157116	AX157116 Sequence
17		13.4	67.0	51	6	AX161711	AX161711 Sequence
18		13.2	66.0	20	6	AX195334	AX195334 Sequence
19	c	13.2	66.0	33	6	A44871	A44871 Sequence 12
20		13.2	66.0	42	6	E03621	E03621 DNA primer
21	c	13.2	66.0	50	6	AX165808	AX165808 Sequence
22	c	13.2	66.0	95	6	AX119976	AX119976 Sequence
23	c	13.2	66.0	95	6	AX138025	AX138025 Sequence
24	c	13.2	66.0	99	9	AB011248	AB011248 Homo sapi
25		13.2	66.0	100	1	CE1TN3IS2	M25449 Plasmid Col
26	c	12.8	64.0	18	6	AX481122	AX481122 Sequence
27	c	12.8	64.0	20	6	E22417	E22417 Antisense n
28		12.8	64.0	26	6	A88149	A88149 Sequence 29
29		12.8	64.0	26	6	A90116	A90116 Sequence 29
30	c	12.8	64.0	42	6	E22401	E22401 Antisense n
31		12.8	64.0	48	6	AR193496	AR193496 Sequence
32		12.8	64.0	51	6	AX160975	AX160975 Sequence
33	c	12.8	64.0	54	6	AR042317	AR042317 Sequence
34	c	12.8	64.0	55	6	AR194774	AR194774 Sequence
35	c	12.8	64.0	61	6	A89690	A89690 Sequence 3
36	c	12.8	64.0	72	1	FPLMREPE	M63628 Escherichia
37	c	12.8	64.0	81	14	AF227771	AF227771 Hepatitis
38	c	12.8	64.0	81	14	AF227773	AF227773 Hepatitis
39	c	12.8	64.0	81	14	AF227774	AF227774 Hepatitis
40	c	12.8	64.0	81	14	AF227775	AF227775 Hepatitis
41	c	12.8	64.0	81	14	AF227776	AF227776 Hepatitis
42	c	12.8	64.0	81	14	AF227777	AF227777 Hepatitis
43	c	12.8	64.0	81	14	AF227778	AF227778 Hepatitis
44	c	12.8	64.0	81	14	AF227779	AF227779 Hepatitis
45	c	12.8	64.0	81	14	AF227780	AF227780 Hepatitis

ALIGNMENTS

RESULT 1
AX286668
LOCUS AX286668 57 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 9 from Patent WO0181554.
ACCESSION AX286668
VERSION AX286668.1 GI:17048736
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Conzelmann,K.K.
TITLE Pneumovirus ns proteins antagonising the interferon (ifn) response
JOURNAL Patent: WO 0181554-A 9 01-NOV-2001;
Conzelmann, Karl-Klaus, Prof. Dr. (DE)

```
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer hNS1-NotI5"
BASE COUNT       14 a     16 c     12 g     15 t
ORIGIN

  Query Match             77.0%;  Score 15.4;  DB 6;  Length 57;
  Best Local Similarity   94.1%;  Pred. No. 2e+04;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 AGCGGCAGCCCCCTCTC 18
          ||||||| ||||||||||
Db      6 AGCGGCCGCCCCCTCTC 22


RESULT 2
A00424/c
LOCUS       A00424                    60 bp    DNA     linear   PAT 28-JAN-1993
DEFINITION  Nucleotide sequence 11 from patent number WO8903877.
ACCESSION   A00424
VERSION     A00424.1  GI:14496
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       11 a     12 c     26 g     11 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 6;  Length 60;
  Best Local Similarity   88.9%;  Pred. No. 3.6e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 AGCGGCAGCCCCCTCTCC 19
          ||| ||||||||||| |
Db     22 AGCCGCAGCCCCCTCTAC 5


RESULT 3
CHKER5LD/c
LOCUS       CHKER5LD                  80 bp    DNA     linear   VRT 28-APR-1993
DEFINITION  Chicken RAV-O endogenous retrovirus, 5' leader region.
ACCESSION   M13372
VERSION     M13372.1  GI:211747
KEYWORDS    mutational analysis.
SOURCE      Chicken DNA, clone pLD6E2.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 80)
  AUTHORS   Katz,R.A., Terry,R.W. and Skalka,A.M.
  TITLE     A conserved cis-acting sequence in the 5' leader of avian sarcoma
            virus RNA is required for packaging
  JOURNAL   J. Virol. 59 (1), 163-167 (1986)
  MEDLINE   86227708
   PUBMED   3012114
COMMENT     Draft entry and clean copy of sequence for [1] kindly provided by
            R.A.Katz, 10-OCT-1986.
            The 5' leader of the RAV-O endogenous retrovirus was inserted into
            Rous sarcoma virus (SR-B).  Deletions (positions 17-47) in the 5'
            leader region of the insert resulted in loss of infectivity of
            SR-B, although the envelope gene was expressed normally.  The
            amount of mutant viral RNA encapsidated into virions was severely
            reduced, despite the presence of helper-virus.  The deleted DNA is
            probably an essential cis-acting packaging signal in SR-B.
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
     mRNA            <1. .>80
                     /product="env mRNA"
BASE COUNT       17 a     18 c     35 g     10 t
ORIGIN      59 bp upstream of SacI site.

  Query Match             74.0%;  Score 14.8;  DB 5;  Length 80;
  Best Local Similarity   88.9%;  Pred. No. 3.5e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 AGCGGCAGCCCCCTCTCC 19
          ||| ||||||||||| |
Db     23 AGCCGCAGCCCCCTCTAC 6


RESULT 4
AX073623
LOCUS       AX073623                  55 bp    DNA     linear   PAT 06-FEB-2001
DEFINITION  Sequence 14 from Patent WO0104335.
ACCESSION   AX073623
VERSION     AX073623.1  GI:12710046
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Buchholz,U., Collins,P.L., Murphy,B.R., Whitehead,S.S. and
            Krempl,C.D.
  TITLE     Human-bovine chimeric respiratory syncytial virus vaccines
  JOURNAL   Patent: WO 0104335-A 14 18-JAN-2001;
            THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic respiratory syncytial virus"
BASE COUNT       14 a     16 c     11 g     14 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 6;  Length 55;
  Best Local Similarity   93.8%;  Pred. No. 5.5e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GCGGCAGCCCCCTCTC 18
          ||||| ||||||||||
Db     22 GCGGCCGCCCCCTCTC 37


RESULT 5
AX359862
LOCUS       AX359862                  55 bp    DNA     linear   PAT 13-FEB-2002
DEFINITION  Sequence 13 from Patent WO0200693.
ACCESSION   AX359862
VERSION     AX359862.1  GI:18675538
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Krempl,C.D., Collins,P.L., Murphy,B.R., Buchholz,U. and
            Whitehead,S.S.
  TITLE     Respiratory syncytial virus vaccines expressing protective antigens
            from promotor-proximal genes
  JOURNAL   Patent: WO 0200693-A 13 03-JAN-2002;
            THE DEPARTMENT OF HEALTH & HUMAN SERVICES (US)
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Artificial Respiratory Syncytial Virus"
BASE COUNT       14 a     16 c     11 g     14 t
ORIGIN
```

```
  Query Match            72.0%;  Score 14.4;  DB 6;  Length 55;
  Best Local Similarity  93.8%;  Pred. No. 5.5e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GCGGCAGCCCCCTCTC 18
          ||||| ||||||||||
Db     22 GCGGCCGCCCCCTCTC 37


RESULT 6
AX286671
LOCUS       AX286671                  60 bp    DNA     linear   PAT 21-NOV-2001
DEFINITION  Sequence 12 from Patent WO0181554.
ACCESSION   AX286671
VERSION     AX286671.1  GI:17048739
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Conzelmann,K.K.
  TITLE     Pneumovirus ns proteins antagonising the interferon (ifn) response
  JOURNAL   Patent: WO 0181554-A 12 01-NOV-2001;
            Conzelmann, Karl-Klaus, Prof. Dr. (DE)
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer mNS1-NotIEcoRV5"
BASE COUNT       13 a     15 c     14 g     18 t
ORIGIN

  Query Match            72.0%;  Score 14.4;  DB 6;  Length 60;
  Best Local Similarity  93.8%;  Pred. No. 5.5e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GCGGCAGCCCCCTCTC 18
          ||||| ||||||||||
Db     10 GCGGCCGCCCCCTCTC 25


RESULT 7
AX204378/c
LOCUS       AX204378                  51 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 484 from Patent WO0148245.
ACCESSION   AX204378
VERSION     AX204378.1  GI:15393911
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0148245-A 484 05-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       26
                     /note="single nucleotide polymorphism
                     Accession number cg40310734"
BASE COUNT       3 a     15 c     20 g     13 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 6;  Length 51;
  Best Local Similarity  84.2%;  Pred. No. 6.9e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCC 19
          |||| ||||| ||| ||||
Db     19 GAGCAGCAGCACCCACTCC 1


RESULT 8
AX481058/c
LOCUS       AX481058                  18 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 118 from Patent WO0246412.
ACCESSION   AX481058
VERSION     AX481058.1  GI:22217610
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Rebar,E., Jamieson,A., Liu,Q., Liu,P.Q., Wolffe,A., Eisenberg,S.P.
            and Jarvis,E.
  TITLE     Regulation of angiogenesis with zinc finger proteins
  JOURNAL   Patent: WO 0246412-A 118 13-JUN-2002;
            Sangamo Biosciences Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="target"
BASE COUNT       3 a      2 c     11 g      1 t      1 others
ORIGIN

  Query Match            70.0%;  Score 14;  DB 6;  Length 18;
  Best Local Similarity  87.5%;  Pred. No. 9.7e+04;
  Matches   14;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 CGGCAGCCCCCTCTCC 19
          | ||:||||||||||||
Db     16 CTGCRGCCCCCTCTCC 1


RESULT 9
AR194772/c
LOCUS       AR194772                  55 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 16 from patent US 6348596.
ACCESSION   AR194772
VERSION     AR194772.1  GI:20241364
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Lee,L.G., Graham,R.J., Mullah,K.B. and Haxo,F.T.
  TITLE     Non-fluorescent asymmetric cyanine dye compounds useful for
            quenching reporter dyes
  JOURNAL   Patent: US 6348596-A 16 19-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="unknown"
BASE COUNT       5 a     25 c     15 g     10 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 55;
  Best Local Similarity  80.0%;  Pred. No. 1.3e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCCA 20
          ||  ||||||||||  ||||
Db     52 GATGGGCAGCCCCCAGTCCA 33


RESULT 10
P531/c
LOCUS       P531                      100 bp    DNA     linear   SYN 31-OCT-2001
DEFINITION  Plasmid pMM5-3, with DNA from pMB8, pHS1 and Tn3, pHS1/Tn3 (IR-L)
            junction.
```

```
ACCESSION   K02977
VERSION     K02977.1  GI:150478
KEYWORDS    recombination joint; transposon.
SEGMENT     1 of 4
SOURCE      Recombinant plasmid pMM5-3.
  ORGANISM  Recombinant plasmid pMM5-3
            artificial sequences; vectors.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   McCormick,M., Wishart,W., Ohtsubo,H., Heffron,F. and Ohtsubo,E.
  TITLE     Plasmid cointegrates and their resolution mediated by transposon
            Tn3 mutants
  JOURNAL   Gene 15 (2-3), 103-118 (1981)
  MEDLINE   82051303
   PUBMED   6271635
COMMENT     Recombinant plasmids were generated in vivo, between a
            temperature-sensitive replication mutant of pSC101 (pSC1), which
            codes for tetracycline resistance, and pMB8, carrying the
            ampicillin transposon Tn3, When the repressor gene of Tn3 was
            inactive, all cells examined contained cointegrate recombinations.
            Complete Tn3 mutant #5 DNA was present at both junctions between
            the parental plasmids, in a direct orientation and a 5 bp
            (positions 53-57) pHS1 sequence was directly duplicated at the
            pHS1/Tn3 #5 junctions in pMM5-3.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Recombinant plasmid pMM5-3"
                     /db_xref="taxon:2602"
                     /note="obtained from in vivo recombination of clones pMB8
                     (containing Tn3 mutant #5) and pHS1 (temperature-sensitive
                     mutant of pSC101)"
BASE COUNT       22 a     21 c     32 g     20 t      5 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 12;  Length 100;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCCA 20
          ||||| ||| ||||||  ||
Db     71 GAGCGTCAGACCCCTCAACA 52


RESULT 11
AX248931
LOCUS       AX248931                  31 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 1010 from Patent WO0166800.
ACCESSION   AX248931
VERSION     AX248931.1  GI:15863554
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Cargill,M., Ireland,J.S. and Lander,E.S.
  TITLE     Human single nucleotide polymorphisms
  JOURNAL   Patent: WO 0166800-A 1010 13-SEP-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        9 a     12 c      7 g      2 t      1 others
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 31;
  Best Local Similarity   82.4%;  Pred. No. 1.7e+05;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CGGCAGCCCCCTCTCCA 20
          | ||||||||:|| |||
Db      6 CAGCAGCCCCMTCGCCA 22


RESULT 12
AX207259/c
LOCUS       AX207259                  37 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 6 from Patent WO0155366.
ACCESSION   AX207259
VERSION     AX207259.1  GI:15395042
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Schumacher,A.N. and Kessels,H.W.
  TITLE     T cell receptor libraries
  JOURNAL   Patent: WO 0155366-A 6 02-AUG-2001;
            Vereniging Het Nederlands Kanker Instituut (NL)
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer F5 alfa-bottom"
BASE COUNT        8 a     12 c     14 g      3 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 37;
  Best Local Similarity   93.3%;  Pred. No. 1.6e+05;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCT 15
          ||||||| |||||||
Db     18 GAGCGGCCGCCCCCT 4


RESULT 13
AX207261/c
LOCUS       AX207261                  42 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 8 from Patent WO0155366.
ACCESSION   AX207261
VERSION     AX207261.1  GI:15395044
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Schumacher,A.N. and Kessels,H.W.
  TITLE     T cell receptor libraries
  JOURNAL   Patent: WO 0155366-A 8 02-AUG-2001;
            Vereniging Het Nederlands Kanker Instituut (NL)
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer F5 beta-bottom"
BASE COUNT        7 a      8 c     15 g     12 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 42;
  Best Local Similarity   93.3%;  Pred. No. 1.6e+05;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCT 15
          ||||||| |||||||
Db     18 GAGCGGCCGCCCCCT 4


RESULT 14
AX161712
LOCUS       AX161712                  50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5040 from Patent WO0140521.
ACCESSION   AX161712
VERSION     AX161712.1  GI:14543043
KEYWORDS    .
SOURCE      human.
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5040 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg43981323"
     misc_feature    26
                     /note="2 of 2 allelic variants (5039 is other entry)"
BASE COUNT        3 a     21 c     18 g      8 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 50;
  Best Local Similarity   93.3%;  Pred. No. 1.6e+05;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 CGGCAGCCCCCTCTC 18
           ||||| |||||||||
Db      34 CGGCATCCCCCTCTC 48


RESULT 15
AX157115/c
LOCUS       AX157115                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 443 from Patent WO0140521.
ACCESSION   AX157115
VERSION     AX157115.1  GI:14538446
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 443 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (444 is other entry)
                     Accession number cg44925402"
BASE COUNT       11 a     11 c     17 g     12 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 51;
  Best Local Similarity   93.3%;  Pred. No. 1.6e+05;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCCT 15
           |||| ||||||||||
Db      43 GAGCCGCAGCCCCCT 29


Search completed: December  3, 2002, 18:11:52
Job time : 310.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-1
Perfect score: 20
Sequence: 1 gagcggcagccccctctcca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31431	Human neuropilin m
	2	14.4	72.0	50	22	AAL33643	Human SNP oligonuc
	3	14.4	72.0	55	22	AAF25148	Nucleotide sequenc
	4	14.4	72.0	55	24	ABK13065	Recombinant BRSV g
	5	14.2	71.0	30	16	AAQ97261	Not-dC oligo prime
c	6	14.2	71.0	51	22	AAH79869	Human DNA containi
c	7	14	70.0	18	24	ABT04101	Human VEGF zinc fi
c	8	13.8	69.0	65	24	ABN52872	Mouse spliced tran
	9	13.8	69.0	90	14	AAQ49420	Cytochrome P450 ho
c	10	13.6	68.0	55	24	ABA91984	Single nucleotide
c	11	13.6	68.0	57	20	ABA96211	Extracellular sign
	12	13.6	68.0	65	24	ABN57612	Mouse spliced tran
c	13	13.6	68.0	84	20	AAX87463	Plasmid pmtrx PCR
c	14	13.6	68.0	92	21	AAC07582	Human secreted pro
c	15	13.4	67.0	37	22	AAH43131	Primer: F5-alpha-b
c	16	13.4	67.0	42	22	AAH43133	Primer: F5-beta-bo
	17	13.4	67.0	50	22	AAI78099	Human silent SNP c
c	18	13.4	67.0	51	22	AAI73502	Human silent SNP c
c	19	13.4	67.0	51	22	AAI73503	Human silent SNP c
	20	13.4	67.0	51	22	AAI78098	Human silent SNP c
	21	13.2	66.0	20	22	AAH48586	Human fascin assoc
	22	13.2	66.0	27	19	AAV21278	Tissue plasminogen
	23	13.2	66.0	30	17	AAT09304	Murine anti-Protei
	24	13.2	66.0	31	21	AAA51765	31-mer CYP3A5 vari
	25	13.2	66.0	31	21	AAA51766	31-mer CYP3A5 wild
c	26	13.2	66.0	36	22	AAI65827	Nucleotide sequenc
	27	13.2	66.0	42	13	AAQ25735	N-terminal extrace
c	28	13.2	66.0	50	23	ABL01012	Human SNP involvin
c	29	13.2	66.0	95	24	ABK11008	Oligonucleotide MN
c	30	13	65.0	36	22	AAF83314	Human neuropilin-1
c	31	13	65.0	45	13	AAQ32238	Human NT-4 probe.
c	32	13	65.0	45	15	AAQ54734	Rat NT-4 DNA probe
	33	13	65.0	93	21	AAC18940	Human secreted pro
c	34	12.8	64.0	18	24	ABT04114	Rat VEGF zinc fing
	35	12.8	64.0	20	20	AAX15781	Antisense oligonuc
c	36	12.8	64.0	20	20	AAX15614	Fragment of upstre
	37	12.8	64.0	26	19	AAV48708	junB gene antisens
c	38	12.8	64.0	27	22	AAH26554	Kidney cancer spec
c	39	12.8	64.0	31	22	AAI30775	Human single nucle
c	40	12.8	64.0	42	20	AAX15598	Fragment of upstre
	41	12.8	64.0	47	21	AAZ68973	Human map-related
	42	12.8	64.0	48	21	AAA93874	PCR primer used to
	43	12.8	64.0	49	20	AAV72183	Human NBC PCR prim
	44	12.8	64.0	50	22	AAL29219	Human SNP oligonuc
	45	12.8	64.0	51	22	AAI77362	Human silent SNP c

ALIGNMENTS

RESULT 1
AAZ31431
ID AAZ31431 standard; DNA; 20 BP.
XX
AC AAZ31431;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3601.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 16; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 3 A; 10 C; 5 G; 2 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGCGGCAGCCCCCTCTCCA 20
||||||||||||||||||||
Db 1 GAGCGGCAGCCCCCTCTCCA 20

RESULT 2
AAL33643
ID AAL33643 standard; DNA; 50 BP.
XX
AC AAL33643;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #6851.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 3341; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 50 BP; 10 A; 16 C; 18 G; 6 T; 0 other;

Query Match 72.0%; Score 14.4; DB 22; Length 50;
Best Local Similarity 93.8%; Pred. No. 5.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 GGCAGCCCCCTCTCCA 20
||||||||| ||||||
Db 17 GGCAGCCCCATCTCCA 32

RESULT 3
AAF25148
ID AAF25148 standard; DNA; 55 BP.
XX
AC AAF25148;
XX
DT 30-APR-2001 (first entry)
XX
DE Nucleotide sequence of a recombinant Respiratory syncytial virus.
XX
KW Respiratory syncytial virus; RSV; attenuated vaccine; RSV A; RSV B;
KW BRSV; HSRV; ss.
XX
OS Respiratory syncytial virus.
XX
PN WO200104335-A2.
XX
PD 18-JAN-2001.
XX
PF 24-JUN-2000; 2000WO-US17755.
XX
PR 09-JUL-1999; 99US-0143132.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Buchholz U, Collins PL, Murphy BR, Whitehead SS, Krempl CD;
XX
DR WPI; 2001-103088/11.
XX
PT Isolated chimeric human-bovine respiratory syncytial virus (RSV),
PT useful in an attenuated vaccine to elicits an immune response against
PT either or both human RSV A or RSV B -
XX
PS Disclosure; Page 145; 148pp; English.
XX
CC The specification describes a chimeric human-bovine respiratory syncytial
CC virus (RSV) that is infectious and attenuated in humans. The virus
CC comprises a major nucleocapsid protein, a nucleocapsid phosphoprotein,
CC a large polymerase protein, a RNA polymerase elongation factor, and a
CC partial or complete RSV background genome, or antigenome of a human RSV
CC or bovine RSV, combined with one or more heterologous genes or genome
CC segment of a different RSV to form a human-bovine chimeric RSV genome
CC or antigenome. The chimeric RSV is useful in an attenuated vaccine to
CC elicits an immune response against either or both human RSV A or RSV B.
CC The present sequence represents a fragment of a recombinant RSV.
XX
SQ Sequence 55 BP; 14 A; 16 C; 11 G; 14 T; 0 other;

```
  Query Match            72.0%;  Score 14.4;  DB 22;  Length 55;
  Best Local Similarity  93.8%;  Pred. No. 5.2e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GCGGCAGCCCCCTCTC 18
          ||||| ||||||||||
Db     22 GCGGCCGCCCCCTCTC 37


RESULT 4
ABK13065
ID   ABK13065 standard; DNA; 55 BP.
XX
AC   ABK13065;
XX
DT   23-APR-2002  (first entry)
XX
DE   Recombinant BRSV genome construct rBRSV/A2-G1F2 leader sequence.
XX
KW   Respiratory syncytial virus; RSV; vaccine; immunostimulatory;
KW   antiviral; gene therapy; rBRSV/A2-G1F2; ds.
XX
OS   Bovine respiratory syncytial virus.
OS   Synthetic.
XX
PN   WO200200693-A2.
XX
PD   03-JAN-2002.
XX
PF   22-JUN-2001; 2001WO-US20107.
XX
PR   23-JUN-2000; 2000US-213708P.
XX
PA   (USGO ) US GOVERNMENT.
XX
PI   Krempl CD,  Collins PL,  Murphy BR,  Buchholz U,  Whitehead SS;
XX
DR   WPI; 2002-090518/12.
XX
PT   An isolated infectious recombinant respiratory syncytial virus (RSV)
PT   having one or more shifted RSV gene(s) or genome segment(s) within the
PT   recombinant genome or antigenome, useful as an attenuated vaccine
PT   against RSV strains -
XX
PS   Example 3; Fig 6; 168pp; English.
XX
CC   The invention relates to an isolated infectious recombinant respiratory
CC   syncytial virus (RSV) having one or more shifted RSV gene(s) or genome
CC   segment(s) within the recombinant genome or antigenome that is/are
CC   positionally shifted to a more promoter-proximal or promoter-distal
CC   position relative to a position of the RSV gene(s) or genome segment(s)
CC   within a wild type RSV genome or antigenome. Also described is (1) a
CC   method for stimulating the immune system of an individual to induce
CC   protection against RSV which comprises administering to the individual an
CC   immunologically sufficient amount of the recombinant RSV combined with a
CC   physiologically acceptable carrier; and (2) an isolated infectious
CC   chimeric RSV comprising a major nucleocapsid protein, a nucleocapsid
CC   phosphoprotein, a large polymerase protein, a RNA polymerase elongation
CC   factor, and a partial or complete bovine RSV background genome or
CC   antigenome combined with heterologous gene(s) and/or genome segment(s)
CC   of a human RSV selected from heterologous gene(s) and/or genome
CC   segment(s) of RSV NS1, NS2, M, SH, G, and/or F, to form a human-bovine
CC   chimeric RSV genome or antigenome. The recombinant RSV is useful in an
CC   attenuated vaccine to elicit an immune response against either human
CC   RSV A or RSV B or both human RSV A and RSV B. The present sequence
CC   represents recombinant BRSV genome construct rBRSV/A2-G1F2 leader
CC   sequence as described in the method of the invention.
XX
SQ   Sequence 55 BP; 14 A; 16 C; 11 G; 14 T; 0 other;

  Query Match            72.0%;  Score 14.4;  DB 24;  Length 55;
  Best Local Similarity  93.8%;  Pred. No. 5.2e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GCGGCAGCCCCCTCTC 18
          ||||| ||||||||||
Db     22 GCGGCCGCCCCCTCTC 37


RESULT 5
AAQ97261
ID   AAQ97261 standard; cDNA; 30 BP.
XX
AC   AAQ97261;
XX
DT   28-DEC-1995  (first entry)
XX
DE   Not-dC oligo primer.
XX
KW   Astrocyte-type 1; AT1; neuron; ss.
XX
OS   Synthetic.
XX
PN   WO9517203-A1.
XX
PD   29-JUN-1995.
XX
PF   22-DEC-1994;   94WO-US14771.
XX
PR   20-DEC-1994;   94US-0359480.
PR   22-DEC-1993;   93US-0172327.
PR   18-JUL-1994;   94US-0275709.
XX
PA   (UYNE-) UNIV NEW JERSEY.
XX
PI   Black IB,  Dreyfuss CF,  Schaar DG;
XX
DR   WPI; 1995-240472/31.
XX
PT   New astrocyte-derived neurotrophic factor proteins - related nucleic
PT   acid, vectors and transformed cells, useful for stimulating neuronal
PT   cell survival and growth
XX
PS   Example 2; Page 43; 107pp; English.
XX
CC   cDNA was synthesized from astrocyte total RNA using reverse
CC   transcriptase primed by a Not-dt oligo (given in AAQ97260).
CC   Products were used for non-sequence specific cDNA amplification,
CC   using Not-dT (AAQ97260) and Not-dC, asymmetric PCR and subtractive
CC   hybridization to isolate novel astrocyte type 1 cDNA clones
CC   (AAQ97243-46).
XX
SQ   Sequence 30 BP; 1 A; 21 C; 7 G; 1 T; 0 other;

  Query Match            71.0%;  Score 14.2;  DB 16;  Length 30;
  Best Local Similarity  84.2%;  Pred. No. 6.4e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCC 19
          ||||||| |||||| | ||
Db      8 GAGCGGCCGCCCCCCCCCC 26


RESULT 6
AAH79869/c
ID   AAH79869 standard; DNA; 51 BP.
XX
AC   AAH79869;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human DNA containing single nucleotide polymorphism SEQ ID NO. 484.
XX
KW   Human; single nucleotide polymorphism; SNP; angiopoietin;
KW   4-hydroxybutyrate; dehydrogenase; protein therapy;
KW   adenosine triphosphate-dependent RNA helicase;
```

KW major histocompatibility complex Class I histocompatibility antigen; MHC;
KW phosphoglycerate kinase; immunosuppressive; immunostimulatory;
KW antirheumatic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;
KW antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200148245-A2.
XX
PD 05-JUL-2001.
XX
PF 27-DEC-2000; 2000WO-US35346.
XX
PR 27-DEC-1999; 99US-0472688.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-418297/44.
XX
PT Polymorphic nucleic acids encoding e.g. angiopoietin, dehydrogenase,
PT adenosine triphosphate-dependent RNA helicase and/or phosphoglycerate
PT kinase, useful for diagnosing and treating, e.g. cancer, autoimmune
PT diseases and infections -
XX
PS Claim 1; Page 195; 484pp; English.
XX
CC The invention relates to nucleic acids (AAH79386-AAH80036) encoding
CC polymorphic variants of proteins (AAG98010-AAG98238) related to
CC angiopoietin, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate
CC (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)
CC Class I histocompatibility antigen and/or phosphoglycerate kinase. These
CC nucleic acid single nucleotide polymorphisms (SNPs) and the encoded
CC proteins have potential immunosuppressive, immunostimulatory,
CC antirheumatic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,
CC antileukemic, neuroprotective and antimicrobial activity and may be
CC useful in gene/protein therapy, vaccines, modulation of the expression
CC and activity of proteins related to angiopoietin, 4-hydroxybutyrate,
CC dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase,
CC major histocompatibility complex (MHC) Class I histocompatibility antigen
CC and/or phosphoglycerate kinase. Disorders that may be prevented,
CC diagnosed and/or treated by the above methods include multifactorial
CC diseases with a genetic component, such as autoimmune diseases (e.g.
CC rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus
CC erytheromatosus and Grave's disease), inflammation, cancer (e.g. cancers
CC of the bladder, brain, breast, colon and kidney, leukemia), diseases of
CC the nervous system, an infection of pathogenic organisms. They may also
CC be used to alter phenotypic traits such as longevity, appearance,
CC strength, speed and endurance.
XX
SQ Sequence 51 BP; 3 A; 15 C; 20 G; 13 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%; Pred. No. 6.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAGCGGCAGCCCCCTCTCC 19
|||| ||||| ||| ||||
Db 19 GAGCAGCAGCACCCACTCC 1

RESULT 7
ABT04101/c
ID ABT04101 standard; DNA; 18 BP.
XX
AC ABT04101;
XX
DT 25-SEP-2002 (first entry)
XX
DE Human VEGF zinc finger target site SEQ ID NO: 118.
XX
KW Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnerary;
KW antiulcer; cytostatic; antipsoriatic; antidiabetic; ophthalmological;
KW osteopathic; antiinfertility; ds.
XX
OS Homo sapiens.
XX
PN WO200246412-A2.
XX
PD 13-JUN-2002.
XX
PF 06-DEC-2001; 2001WO-US46861.
XX
PR 07-DEC-2000; 2000US-0733604.
PR 12-DEC-2000; 2000US-0736083.
PR 30-APR-2001; 2001US-0846033.
XX
PA (SANG-) SANGAMO BIOSCIENCES INC.
XX
PI Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
PI Jarvis E;
XX
DR WPI; 2002-527918/56.
XX
PT New zinc finger protein that binds to target site in vascular
PT endothelial growth factor gene, useful for modulating expression of the
PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
PT ulcer -
XX
PS Claim 1; Page 103; 195pp; English.
XX
CC The present invention relates to a zinc finger protein that binds to a
CC target site in one or more vascular endothelial growth factor (VEGF)
CC genes. The protein is useful for modulating expression of a VEGF gene,
CC thereby regulating angiogenesis and vasculogenesis. This can be used to
CC treat atherosclerosis, ischaemia, arthritis, wounds, ulcers, tumours,
CC diabetic retinopathy or psoriasis. The present sequence is a target DNA
CC sequence for a zinc finger protein shown in the invention.
XX
SQ Sequence 18 BP; 3 A; 2 C; 11 G; 1 T; 1 other;

Query Match 70.0%; Score 14; DB 24; Length 18;
Best Local Similarity 87.5%; Pred. No. 7.8e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CGGCAGCCCCCTCTCC 19
| ||:|||||||||||
Db 16 CTGCRGCCCCCTCTCC 1

RESULT 8
ABN52872/c
ID ABN52872 standard; DNA; 65 BP.
XX
AC ABN52872;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:25620.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.

```
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 25620; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 11 A; 20 C; 22 G; 12 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 65;
  Best Local Similarity   88.2%;  Pred. No. 9.2e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 AGCGGCAGCCCCCTCTC 18
           |||||||||| |||| |
Db      18 AGCGGCAGCCGCCTCCC 2


RESULT 9
AAQ49420
ID   AAQ49420 standard; cDNA; 90 BP.
XX
AC   AAQ49420;
XX
DT   27-APR-1994  (first entry)
XX
DE   Cytochrome P450 homologue cDNA insert pCGP163 start.
XX
KW   Transgenic plants; altered petal colour; partial sequence; ss.
XX
OS   Petunia cv. OGB (Old Glory Blue).
XX
PN   WO9320206-A.
XX
PD   14-OCT-1993.
XX
PF   25-MAR-1993;   93WO-AU00127.
XX
PR   27-MAR-1992;   92AU-0001538.
PR   07-JAN-1993;   93AU-0006698.
XX
PA   (ITFL-) INT FLOWER DEV PTY LTD.
XX
PI   Cornish EC,  Holton TA,  Tanaka Y;
XX
DR   WPI; 1993-336914/42.
XX
PT   Nucleic acid isolate encoding flavonoid-3'-hydroxylase - is used to
PT   create transgenic plants with altered petal colour
XX
PS   Disclosure; Fig 4A; 86pp; English.
XX
CC   The sequence is that of the start of a partial nucleotide sequence
CC   of a cytochrome P450 homologue cDNA insert isolated from pCGP163.
CC   The end of this sequence is given in AAQ58357.
XX
SQ   Sequence 90 BP; 28 A; 22 C; 14 G; 26 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 14;  Length 90;
  Best Local Similarity   88.2%;  Pred. No. 9.1e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GCGGCAGCCCCCTCTCC 19
           || ||| ||||||||||
Db      36 GCTGCATCCCCCTCTCC 52


RESULT 10
ABA91984/c
ID   ABA91984 standard; DNA; 55 BP.
XX
AC   ABA91984;
XX
DT   23-MAY-2002  (first entry)
XX
DE   Single nucleotide polymorphism BAK target oligonucleotide F2662T.
XX
KW   Single nucleotide polymorphism; SNP; detection; Taqman; assay;
KW   quencher; hybridisation; human; probe; ss.
XX
OS   Homo sapiens.
XX
PN   US6348596-B1.
XX
PD   19-FEB-2002.
XX
PF   20-JUL-1999;   99US-0357740.
XX
PR   23-JAN-1998;   98US-0012525.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Lee LG,  Graham RJ,  Mullah KB,  Haxo FT;
XX
DR   WPI; 2002-225175/28.
XX
PT   New non-fluorescent asymmetric cyanide dye compounds, useful for
PT   quenching reporter dyes in nucleic acid hybridisation assays employing
PT   fluorescence energy transfer as means of detection -
XX
PS   Example 4; Column 66; 62pp; English.
XX
CC   The present sequence is that of single nucleotide polymorphism
CC   (SNP) BAK target oligonucleotide F2622T.  This target sequence was
CC   used in a multiplex endopoint single nucleotide polymorphism (SNP)
CC   analysis designed to demonstrate the use of a novel non-fluorescent
CC   asymmetric cyanide dye, nitrothiazole blue (NTB), as an quenching
CC   reporter dye.  A 7-colour homogeneous detection of multiple PCR
CC   products was performed.  The test system was a set of 3 SNPs,
CC   denoted MPO, BAK and LIG.  For each SNP system there were 2
CC   primers (see ABA91969-74) and 2 sequence-specific probes (see
CC   ABA91975-80), each with NTB at the 3' end, and a different reporter
CC   dye (6-FAM, dR110, dR6G, dTMR, DROX and JAZ) at the 5' end.  The 7th
CC   colour was from aluminium phthalocyanine tetrasulfonate, used as a
```

```
CC   passive reference. Following PCR, the reactions were measured on a
CC   luminescence spectrometer in synchronous scanning mode. The
CC   spectral overlap in the set was evaluated by calculation of the
CC   conditioning number of the 7x7 matrix (dye fluorescence versus
CC   wavelength). The small value of the condition number (1.5) proved
CC   that crosstalk between the dyes was minimal. SNP analyses of
CC   known, synthetic target DNA sequences (see ABA91981-90) and genomic
CC   DNA (from human blood samples and Raji (ATCC CCL-86) cells) were
CC   plotted as normalised, subtracted spectra and as data points in dot
CC   plots. The synthetic targets, containing priming sites and both
CC   possible alleles, were constructed by synthesising the forward and
CC   reverse 5' ends of the target with a 10 bp complementary region at
CC   the 3' terminus. The oligonucleotides were combined and the ends
CC   filled in. The BAK target utilised a separate pair of primers for
CC   each allele (see also ABA91985-87). The multiplex PCR system
CC   provides increased sample throughput and potential cost savings.
XX
SQ   Sequence 55 BP; 5 A; 25 C; 15 G; 10 T; 0 other;

  Query Match            68.0%;  Score 13.6;  DB 24;  Length 55;
  Best Local Similarity  80.0%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCCA 20
          ||  ||||||||||  ||||
Db     52 GATGGGCAGCCCCCAGTCCA 33


RESULT 11
ABA96211/c
ID   ABA96211 standard; DNA; 57 BP.
XX
AC   ABA96211;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Extracellular signal-regulated PK-1 related oligonucleotide 5.
XX
KW   Extracellular signal-related protein kinase-1; ss.
XX
OS   Unidentified.
XX
PN   KR98031184-A.
XX
PD   25-JUL-1998.
XX
PF   31-OCT-1996;   96KR-0050731.
XX
PR   31-OCT-1996;   96KR-0050731.
XX
PA   (GLDS ) LG CHEM LTD.
XX
PI   Kim CH,  Cho JM,  Chung HH,  Lee JH;
XX
DR   WPI; 1999-369099/31.
XX
PT   Process for preparing extracellular signal-regulated protein kinase-1
PT   from E. coli  -
XX
PS   Disclosure; Fig 5; 10pp; Korean.
XX
CC   The invention relates to preparing extracellular signal-regulated
CC   protein kinase-1. The present sequence is that of an oligonucleotide,
CC   useful to the invention.
XX
SQ   Sequence 57 BP; 9 A; 13 C; 27 G; 8 T; 0 other;

  Query Match            68.0%;  Score 13.6;  DB 20;  Length 57;
  Best Local Similarity  80.0%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCCA 20
          ||  ||||  |||||  ||  |||
Db     39 GAACGGCTGCCCCTTCACCA 20


RESULT 12
ABN57612
ID   ABN57612 standard; DNA; 65 BP.
XX
AC   ABN57612;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse spliced transcript detection oligonucleotide SEQ ID NO:30360.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Mus musculus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 30360; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 19 A; 18 C; 14 G; 14 T; 0 other;

  Query Match            68.0%;  Score 13.6;  DB 24;  Length 65;
  Best Local Similarity  80.0%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCCA 20
          ||||  ||  ||  ||||||||
```

```
Db     32 GAGCAGCCGCTTCCTCTCCA 51


RESULT 13
AAX87463/c
ID   AAX87463 standard; DNA; 84 BP.
XX
AC   AAX87463;
XX
DT   08-OCT-1999  (first entry)
XX
DE   Plasmid pmtrx PCR primer right.
XX
KW   ET cloning, recE; recT; Escherichia coli; PCR; primer; pmtrx;
KW   mouse; trithorax; homologous recombination; ss.
XX
OS   Synthetic.
OS   Mus sp.
XX
PN   WO9929837-A2.
XX
PD   17-JUN-1999.
XX
PF   07-DEC-1998;   98WO-EP07945.
XX
PR   05-OCT-1998;   98EP-0118756.
PR   05-DEC-1997;   97EP-0121462.
XX
PA   (EMBL-) EMBL EURO LAB MOLEKULARBIOLOGIE.
XX
PI   Buchholz F,  Stewart F,  Zhang Y;
XX
DR   WPI; 1999-457893/38.
XX
PT   A method for cloning DNA using a recE and recT homologous
PT   recombination system
XX
PS   Disclosure; Page 138; 139pp; English.
XX
CC   This is the nucleotide sequence of a primer used in the PCR
CC   amplification of a linear DNA fragment used in a method of the
CC   invention.  The template was pmtrx, a pBluescript vector carrying
CC   mouse trithorax cDNA, and the targeting vector was pZero2.1.  Host
CC   cells and/or vectors expressing Escherichia coli recE and recT
CC   genes (or functionally related genes) are useful in a novel cloning
CC   method, designated ET cloning, which involves homologous
CC   recombination via the recET mechanism (see also AAX87427-29).  The
CC   method covalently combines one preferably linear, extrachromosomal
CC   DNA fragment (the fragment to be cloned), with a second preferably
CC   circular DNA vector.  The method is particularly applicable to
CC   complex or difficult DNA constructs, such as those intended for
CC   homologous recombination in eukaryotic cells.
XX
SQ   Sequence 84 BP; 12 A; 22 C; 30 G; 20 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 20;  Length 84;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCCA 20
          |||||||| ||||| |  ||
Db     81 GAGCGGCACCCCCCACATCA 62


RESULT 14
AAC07582/c
ID   AAC07582 standard; cDNA; 92 BP.
XX
AC   AAC07582;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 11657.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 11657; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC   5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC   used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC   in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors.
XX
SQ   Sequence 92 BP; 17 A; 21 C; 37 G; 17 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 21;  Length 92;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCCA 20
          |||| || ||| |||| |||
Db     34 GAGCCGCCGCCGCCTCGCCA 15


RESULT 15
AAH43131/c
ID   AAH43131 standard; cDNA; 37 BP.
XX
AC   AAH43131;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Primer: F5-alpha-bottom.
XX
KW   T cell receptor; TCR; influenza A; ligand; functional processing;
KW   affinity; polymerase chain reaction; primer; amplify; PCR; ss.
XX
OS   Synthetic.
XX
PN   EP1118661-A1.
XX
PD   25-JUL-2001.
XX
PF   13-JAN-2000; 2000EP-0200110.
XX
```

PR 13-JAN-2000; 2000EP-0200110.
XX
PA (HETN-) HET NEDERLANDS KANKER INST.
XX
PI Schumacher ANM, Kessels HWHG;
XX
DR WPI; 2001-459002/50.
XX
PT Generating receptors with desired specificities for ligands, which upon
PT binding to the ligand under-goes functional processing in order to
PT provide a biological response after ligand-binding -
XX
PS Disclosure; Page 6; 21pp; English.
XX
CC The sequences given in AAH43128-34 are primers which were used in the
CC amplification and cloning of the complementarity determining regions
CC (CDR3) of a T cell receptor which was generated by the method of the
CC invention. The method of the invention allows generation of receptors
CC with desired specificities for ligands, which upon binding to the
CC ligand under-go functional processing in order to provide a biological
CC response after ligand-binding. The method is used for generating
CC libraries of cells with receptors of differing affinities.
XX
SQ Sequence 37 BP; 8 A; 12 C; 14 G; 3 T; 0 other;

Query Match 67.0%; Score 13.4; DB 22; Length 37;
Best Local Similarity 93.3%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 GAGCGGCAGCCCCCT 15
           ||||||| |||||||
Db      18 GAGCGGCCGCCCCCT 4
```

Search completed: November 23, 2002, 06:28:25
Job time : 102.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-1
Perfect score: 20
Sequence: 1 gagcggcagccccctctcca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	14.8	74.0	60	6	5252465-13	Patent No. 5252465
c	2	13.6	68.0	55	4	US-09-357-740-16	Sequence 16, Appl
	3	13.2	66.0	30	5	PCT-US95-07372-3	Sequence 3, Appli
c	4	12.8	64.0	20	3	US-09-120-853-19	Sequence 19, Appl
c	5	12.8	64.0	42	3	US-09-120-853-3	Sequence 3, Appli
	6	12.8	64.0	48	4	US-09-400-653A-49	Sequence 49, Appl
c	7	12.8	64.0	54	1	US-08-311-486C-1107	Sequence 1107, Ap
c	8	12.8	64.0	55	4	US-09-357-740-18	Sequence 18, Appl
	9	12.8	64.0	91	1	US-07-752-101A-53	Sequence 53, Appl
	10	12.8	64.0	93	3	US-08-976-413A-420	Sequence 420, App
	11	12.6	63.0	21	1	US-08-465-590-136	Sequence 136, App
	12	12.6	63.0	21	4	US-08-711-417C-136	Sequence 136, App
	13	12.6	63.0	21	4	US-09-177-650-42	Sequence 42, Appl
	14	12.6	63.0	21	5	PCT-US93-08743-136	Sequence 136, App
	15	12.6	63.0	25	3	US-08-772-512A-18	Sequence 18, Appl
	16	12.6	63.0	25	6	5504194-3	Patent No. 5504194
	17	12.6	63.0	26	4	US-09-271-013-1	Sequence 1, Appli
	18	12.6	63.0	35	1	US-07-832-905B-48	Sequence 48, Appl
	19	12.6	63.0	35	2	US-08-700-757-48	Sequence 48, Appl
	20	12.6	63.0	39	4	US-09-046-158A-9	Sequence 9, Appli
	21	12.6	63.0	41	4	US-09-214-151-5	Sequence 5, Appli
	22	12.6	63.0	42	1	US-07-982-712-4	Sequence 4, Appli
c	23	12.6	63.0	44	1	US-07-832-905B-47	Sequence 47, Appl
c	24	12.6	63.0	44	2	US-08-700-757-47	Sequence 47, Appl
	25	12.6	63.0	69	1	US-07-609-716-29	Sequence 29, Appl
	26	12.6	63.0	69	1	US-08-175-155-27	Sequence 27, Appl
	27	12.6	63.0	69	1	US-08-477-509B-62	Sequence 62, Appl
	28	12.6	63.0	69	3	US-08-482-085B-62	Sequence 62, Appl
	29	12.6	63.0	69	3	US-08-475-411A-29	Sequence 29, Appl
	30	12.6	63.0	69	4	US-08-478-029A-29	Sequence 29, Appl
	31	12.6	63.0	93	2	US-08-592-383-8	Sequence 8, Appli
c	32	12.4	62.0	27	1	US-08-233-030-44	Sequence 44, Appl
c	33	12.4	62.0	65	6	5174993-17	Patent No. 5174993
c	34	12.4	62.0	65	6	5505941-17	Patent No. 5505941
	35	12.4	62.0	69	2	US-08-488-161-102	Sequence 102, App
	36	12.4	62.0	69	3	US-09-273-685-102	Sequence 102, App
	37	12.4	62.0	69	5	PCT-US95-11934-102	Sequence 102, App
c	38	12.2	61.0	21	4	US-09-689-291A-8	Sequence 8, Appli
c	39	12.2	61.0	25	1	US-08-641-627A-17	Sequence 17, Appl
c	40	12.2	61.0	25	1	US-08-641-627A-18	Sequence 18, Appl
c	41	12.2	61.0	25	1	US-08-641-627A-19	Sequence 19, Appl
c	42	12.2	61.0	27	3	US-09-263-023-10	Sequence 10, Appl
c	43	12.2	61.0	27	4	US-09-471-867-10	Sequence 10, Appl
c	44	12.2	61.0	33	2	US-08-750-134A-1	Sequence 1, Appli
c	45	12.2	61.0	33	4	US-09-363-745-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
5252465-13/c
;Patent No. 5252465
; APPLICANT: NIGON, VICTOR-MARC;VERDIER, GERARD;CHEBLOUNE,
;YAHIA;COSSET, FRANCOIS-LOIC;LEGRAS, CATHERINE;REYSS-BRION,
;ASTRID;BELAKEBI, MUSTAPHA;MALLET, FRANCOIS;SAVATIER, PIERE;
;THORAVAL, PIERRICK;SAMARUT, JACQUES;PONCET, DIDIER;BAGNIS,
;CLAUDE;BENCHAIBI, MILOUD
; TITLE OF INVENTION: AVIAN ERYTHROBLASTOSIS VIRUS VECTORS FOR
;INTEGRATION AND EXPRESSION OF HETEROLOGOUS GENES IN AVIAN CELLS
; NUMBER OF SEQUENCES: 14
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/477,833
; FILING DATE: 03-OCT-1988
;SEQ ID NO:13:
; LENGTH:60
5252465-13

Query Match 74.0%; Score 14.8; DB 6; Length 60;
Best Local Similarity 88.9%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      2 AGCGGCAGCCCCCTCTCC 19
          ||| |||||||||||| |
Db     22 AGCCGCAGCCCCCTCTAC 5
```

RESULT 2
US-09-357-740-16/c
; Sequence 16, Application US/09357740
; Patent No. 6348596
; GENERAL INFORMATION:
; APPLICANT: Lee, Linda G.
; APPLICANT: Graham, Ronald J.
; APPLICANT: Mullah, Khairuzzaman B.
; APPLICANT: Haxo, Francis T.
; TITLE OF INVENTION: ASYMMETRIC CYANINE DYE QUENCHERS
; FILE REFERENCE: 9584-007
; CURRENT APPLICATION NUMBER: US/09/357,740
; CURRENT FILING DATE: 1999-07-20
; EARLIER APPLICATION NUMBER: 09/012,525
; EARLIER FILING DATE: 1998-01-23
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 55
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:

```
;   OTHER INFORMATION: Oligonucleotide
US-09-357-740-16

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 55;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCCA 20
          ||  ||||||||||  ||||
Db     52 GATGGGCAGCCCCCAGTCCA 33


RESULT 3
PCT-US95-07372-3
; Sequence 3, Application PC/TUS9507372
;  GENERAL INFORMATION:
;    APPLICANT:  Oklahoma Medical Research Foundation
;    TITLE OF INVENTION:  Calcium Binding Recombinant
;    TITLE OF INVENTION:  Antibody Against Protein C
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET: ' 2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07372
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  OMRF106CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
PCT-US95-07372-3

  Query Match             66.0%;  Score 13.2;  DB 5;  Length 30;
  Best Local Similarity   83.3%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 AGCGGCAGCCCCCTCTCC 19
          |||||| |||||| | ||
Db      2 AGCGGCCGCCCCCCCCCC 19


RESULT 4
US-09-120-853-19/c
; Sequence 19, Application US/09120853
; Patent No. 6057437
; GENERAL INFORMATION:
;  APPLICANT: Kamiya, Kinya
;  APPLICANT: Matsuda, Yoko
;  APPLICANT: Uchida, Kiyoshi
;  TITLE OF INVENTION: AN ANTISENSE NUCLEIC ACID COMPOUND
;  FILE REFERENCE: 07898/030001
;  CURRENT APPLICATION NUMBER: US/09/120,853
;  CURRENT FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: JP 213838/1997
;  EARLIER FILING DATE: 1997-07-25
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Artificial
;   OTHER INFORMATION: nucleic acid sequence
US-09-120-853-19

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 3e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CGGCAGCCCCCTCTCC 19
          | || ||||||||||
Db     20 CTGCGGCCCCCTCTCC 5


RESULT 5
US-09-120-853-3/c
; Sequence 3, Application US/09120853
; Patent No. 6057437
; GENERAL INFORMATION:
;  APPLICANT: Kamiya, Kinya
;  APPLICANT: Matsuda, Yoko
;  APPLICANT: Uchida, Kiyoshi
;  TITLE OF INVENTION: AN ANTISENSE NUCLEIC ACID COMPOUND
;  FILE REFERENCE: 07898/030001
;  CURRENT APPLICATION NUMBER: US/09/120,853
;  CURRENT FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: JP 213838/1997
;  EARLIER FILING DATE: 1997-07-25
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 42
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Artificial
;   OTHER INFORMATION: nucleic acid sequence
US-09-120-853-3

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 42;
  Best Local Similarity   87.5%;  Pred. No. 3e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CGGCAGCCCCCTCTCC 19
          | || ||||||||||
Db     42 CTGCGGCCCCCTCTCC 27


RESULT 6
US-09-400-653A-49
; Sequence 49, Application US/09400653A
; Patent No. 6348311
; GENERAL INFORMATION:
;  APPLICANT:  Kastan, Michael
;  APPLICANT:  Canman, Christine
;  APPLICANT:  Kim, Seong-Tae
;  APPLICANT:  Lim, Dae-Sik
;  APPLICANT:  St. Jude Children's Research Hospital
;  TITLE OF INVENTION:  ATM Kinase Modulation for Screening and Therapies
;  FILE REFERENCE:  2427/1F142
;  CURRENT APPLICATION NUMBER: US/09/400,653A
;  CURRENT FILING DATE:  1999-09-21
```

```
;  PRIOR APPLICATION NUMBER:  09/248,061
;  PRIOR FILING DATE:  1999-02-10
;  NUMBER OF SEQ ID NOS:  49
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 49
;   LENGTH: 48
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR primer
US-09-400-653A-49

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 48;
  Best Local Similarity   87.5%;  Pred. No. 3e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCCTC 16
           ||||||| ||| ||||
Db       7 GAGCGGCCGCCACCTC 22


RESULT 7
US-08-311-486C-1107/c
; Sequence 1107, Application US/08311486C
; Patent No. 5811300
;  GENERAL INFORMATION:
;    APPLICANT:  Sean Sullivan
;    APPLICANT:  Kenneth Draper
;    APPLICANT:  Kevin Kisich
;    APPLICANT:  Dan T. Stinchcomb
;    APPLICANT:  James McSwiggen
;    TITLE OF INVENTION:  RIBOZYME TREATMENT OF
;    TITLE OF INVENTION:  DISEASES OR CONDITIONS
;    TITLE OF INVENTION:  RELATED TO LEVELS OF
;    TITLE OF INVENTION:  TNF-`
;    NUMBER OF SEQUENCES:  1157
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,486C
;      FILING DATE:  September 23, 1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:                  two
;      APPLICATION NUMBER:  08/008,895
;      FILING DATE:  January 19, 1993
;      APPLICATION NUMBER:  07/989,849
;      FILING DATE:  December 7, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  209/166
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  1107:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  54 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-311-486C-1107

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 54;
  Best Local Similarity   87.5%;  Pred. No. 3e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GGCAGCCCCCTCTCCA 20
           || |||||| ||||||
Db      21 GGTAGCCCCTTCTCCA 6


RESULT 8
US-09-357-740-18/c
; Sequence 18, Application US/09357740
; Patent No. 6348596
; GENERAL INFORMATION:
;  APPLICANT: Lee, Linda G.
;  APPLICANT: Graham, Ronald J.
;  APPLICANT: Mullah, Khairuzzaman B.
;  APPLICANT: Haxo, Francis T.
;  TITLE OF INVENTION: ASYMMETRIC CYANINE DYE QUENCHERS
;  FILE REFERENCE: 9584-007
;  CURRENT APPLICATION NUMBER: US/09/357,740
;  CURRENT FILING DATE: 1999-07-20
;  EARLIER APPLICATION NUMBER: 09/012,525
;  EARLIER FILING DATE: 1998-01-23
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 55
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: Oligonucleotide
US-09-357-740-18

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 55;
  Best Local Similarity   87.5%;  Pred. No. 3e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GGCAGCCCCCTCTCCA 20
           ||||||||||  ||||
Db      48 GGCAGCCCCCAGTCCA 33


RESULT 9
US-07-752-101A-53
; Sequence 53, Application US/07752101A
; Patent No. 5326857
;  GENERAL INFORMATION:
;    APPLICANT:  Yamamoto, Fumi-ichiro
;    APPLICANT:  White, Thayer
;    APPLICANT:  Hakomori, Sen-itiroh
;    APPLICANT:  Clausen, Henrik
;    TITLE OF INVENTION:  ABO GENOTYPING
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  U.S.
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/752,101A
```

```
;       FILING DATE:  19910829
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sharkey, Richard G.
;       REGISTRATION NUMBER:  32,629
;       REFERENCE/DOCKET NUMBER:  150036.406C1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  206-622-4900
;       TELEFAX:  206-682-6031
;       TELEX:  3723836
;   INFORMATION FOR SEQ ID NO:  53:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  91 base pairs
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FRAGMENT TYPE:  N-terminal
US-07-752-101A-53

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 91;
  Best Local Similarity   87.5%;  Pred. No. 3e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CGGCAGCCCCCTCTCC 19
           |||||||||| || ||
Db      57 CGGCAGCCCCGTCCCC 72


RESULT 10
US-08-976-413A-420
; Sequence 420, Application US/08976413A
; Patent No. 6127119
;  GENERAL INFORMATION:
;    APPLICANT:  STEPHENS, ANDREW
;    APPLICANT:  GOLD, LARRY
;    APPLICANT:  SPECK, ULRICH
;    TITLE OF INVENTION:  NUCLEIC ACID LIGANDS OF TISSUE TARGET
;    NUMBER OF SEQUENCES:  440
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 diskette, 1.44 MG
;      COMPUTER:  IBM pc compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 8.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/976,413A
;      FILING DATE:  21-NOVEMBER-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/433,124
;      FILING DATE:  03-MAY-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX31/CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  420:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  93
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:  All C's are 2'-F cytosine
;    FEATURE:
;      OTHER INFORMATION:  All U's are 2'-F uracil
US-08-976-413A-420

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 93;
  Best Local Similarity   75.0%;  Pred. No. 3e+03;
  Matches   12;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CGGCAGCCCCCTCTCC 19
           || |||||||:|: |
Db      37 CGACAGCCCCCUCUUC 52


RESULT 11
US-08-465-590-136
; Sequence 136, Application US/08465590
; Patent No. 5824770
;  GENERAL INFORMATION:
;    APPLICANT:  Georgopoulos, Katia A.
;    TITLE OF INVENTION:  IKAROS: A T CELL PATHWAY REGULATORY GENE
;    NUMBER OF SEQUENCES:  164
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 STATE STREET, Suite 510
;      CITY:  BOSTON
;      STATE:  MASSACHUCETTS
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  AscII (text)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,590
;      FILING DATE:  05-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/238,212
;      FILING DATE:  02-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/121,438
;      FILING DATE:  14-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/946,233
;      FILING DATE:  14-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Myers, Paul L.
;      REGISTRATION NUMBER:  35,695
;      REFERENCE/DOCKET NUMBER:  MPG-006C2DV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  136:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-465-590-136

  Query Match             63.0%;  Score 12.6;  DB 1;  Length 21;
```

```
  Best Local Similarity   78.9%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 AGCGGCAGCCCCCTCTCCA 20
           || || | | |||||||||
Db       2 AGGGGAATCTCCCTCTCCA 20


RESULT 12
US-08-711-417C-136
; Sequence 136, Application US/08711417C
; Patent No. 6228611
;   GENERAL INFORMATION:
;        APPLICANT: Georgopoulos, Katia A.
;        TITLE OF INVENTION: IKAROS: A T CELL PATHWAY REGULATORY GENE
;        NUMBER OF SEQUENCES: 202
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows 95
;             SOFTWARE: FastSEQ for Windows Version 2.0b
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/711,417C
;             FILING DATE: 05-Sep-1996
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/238,212
;             FILING DATE: 02-MAY-1994
;             APPLICATION NUMBER: 08/121,438
;             FILING DATE: 14-SEP-1993
;             APPLICATION NUMBER: 07/946,233
;             FILING DATE: 14-SEP-1992
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Myers, Louis P.
;             REGISTRATION NUMBER: 35,965
;             REFERENCE/DOCKET NUMBER: 10287/007001
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617/542-5070
;             TELEFAX: 617/542-8906
;             TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 136:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 21 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 136:
US-08-711-417C-136

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 21;
  Best Local Similarity   78.9%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 AGCGGCAGCCCCCTCTCCA 20
           || || | | |||||||||
Db       2 AGGGGAATCTCCCTCTCCA 20


RESULT 13
US-09-177-650-42
; Sequence 42, Application US/09177650
; Patent No. 6413719
; GENERAL INFORMATION:
;  APPLICANT: Leppert, Mark F.
;  APPLICANT: Singh, Nanda
;  APPLICANT: Charlier, Carole
;  TITLE OF INVENTION: KCNQ2 AND KCNQ3 - POTASSIUM CHANNEL GENES WHICH ARE
;  TITLE OF INVENTION: MUTATED IN BENIGN FAMILIAL NEONATAL CONVULSIONS (BFNC)
;  TITLE OF INVENTION: AND OTHER EPILEPSIES
;  FILE REFERENCE: 2323-134
;  CURRENT APPLICATION NUMBER: US/09/177,650
;  CURRENT FILING DATE: 1998-10-23
;  EARLIER APPLICATION NUMBER: 60/063,147
;  EARLIER FILING DATE: 1997-10-24
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-177-650-42

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 21;
  Best Local Similarity   78.9%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 AGCGGCAGCCCCCTCTCCA 20
           ||||| || ||||||  ||
Db       1 AGCGGGAGGCCCCTCCTCA 19


RESULT 14
PCT-US93-08743-136
; Sequence 136, Application PC/TUS9308743
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  IKAROS: A T CELL PATHWAY REGULATORY GENE
;    NUMBER OF SEQUENCES:  152
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08743
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 946,233
;      FILING DATE:  14-SEP-1992
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  136:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
PCT-US93-08743-136

  Query Match             63.0%;  Score 12.6;  DB 5;  Length 21;
  Best Local Similarity   78.9%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 AGCGGCAGCCCCCTCTCCA 20
           || || | | |||||||||
Db       2 AGGGGAATCTCCCTCTCCA 20


RESULT 15
US-08-772-512A-18
; Sequence 18, Application US/08772512A
; Patent No. 6022705
;  GENERAL INFORMATION:
;    APPLICANT:  Soderlund, David M.
;    APPLICANT:  Knipple, Douglas C.
;    APPLICANT:  Ingles, Patricia J.
;    TITLE OF INVENTION:  INSECT SODIUM CHANNELS FROM
;    TITLE OF INVENTION:  INSECTICIDE-SUSCEPTIBLE AND INSECTICIDE-RESISTANT HOUSE
```

```
;    TITLE OF INVENTION:  FLIES
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  P.O. Box 1051, Clinton Square
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/772,512A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/608,618
;      FILING DATE:  01-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Braman, Susan J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/601(CRFD-1657)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-772-512A-18

  Query Match             63.0%;  Score 12.6;  DB 3;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCCTCTCC 19
           ||||||| |||||  | ||
Db       2 GAGCGGCGGCCCCGGCCCC 20


Search completed: November 23, 2002, 06:35:55
Job time : 23.55 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-1
Perfect score: 20
Sequence: 1 gagcggcagccccctctcca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.4	72.0	55	10	US-09-887-469-13	Sequence 13, Appl
	2	13.4	67.0	31	10	US-09-801-274-1010	Sequence 1010, Ap
c	3	13.2	66.0	69	10	US-09-918-063-35	Sequence 35, Appl
	4	13.2	66.0	69	10	US-09-918-063-37	Sequence 37, Appl
c	5	12.8	64.0	27	10	US-09-783-708-7	Sequence 7, Appli
c	6	12.8	64.0	61	10	US-09-359-672-3	Sequence 3, Appli
	7	12.8	64.0	100	10	US-09-861-893-41	Sequence 41, Appl
c	8	12.6	63.0	27	10	US-09-895-141-40	Sequence 40, Appl
	9	12.6	63.0	41	10	US-09-908-660-5	Sequence 5, Appli
c	10	12.4	62.0	31	10	US-09-801-274-1263	Sequence 1263, Ap
c	11	12.4	62.0	87	10	US-09-920-118-1	Sequence 1, Appli
	12	12.4	62.0	87	10	US-09-920-118-2	Sequence 2, Appli
	13	12.2	61.0	18	12	US-10-037-616-17	Sequence 17, Appl
c	14	12.2	61.0	35	10	US-09-732-348-8	Sequence 8, Appli
c	15	12.2	61.0	38	10	US-09-732-348-20	Sequence 20, Appl
c	16	12.2	61.0	44	10	US-09-916-230-25	Sequence 25, Appl
c	17	12.2	61.0	48	9	US-09-905-291A-355	Sequence 355, App
c	18	12.2	61.0	48	10	US-09-909-320-355	Sequence 355, App
c	19	12.2	61.0	48	10	US-09-909-088B-355	Sequence 355, App
	20	12.2	61.0	56	10	US-09-983-965-4006	Sequence 4006, Ap
c	21	12.2	61.0	83	10	US-09-864-761-26469	Sequence 26469, A
	22	12.2	61.0	84	10	US-09-864-761-23396	Sequence 23396, A
c	23	12	60.0	35	12	US-10-011-125-7	Sequence 7, Appli
	24	12	60.0	36	10	US-09-740-668A-57	Sequence 57, Appl
	25	12	60.0	43	12	US-10-011-125-6	Sequence 6, Appli
	26	12	60.0	46	10	US-09-859-773A-6	Sequence 6, Appli
	27	12	60.0	56	10	US-09-294-093B-19	Sequence 19, Appl
c	28	12	60.0	61	9	US-09-946-807-1302	Sequence 1302, Ap
c	29	12	60.0	61	10	US-09-795-668-1302	Sequence 1302, Ap
c	30	12	60.0	61	10	US-09-795-686-1302	Sequence 1302, Ap
	31	12	60.0	75	10	US-09-864-761-28813	Sequence 28813, A
c	32	12	60.0	78	10	US-09-864-761-24773	Sequence 24773, A
	33	12	60.0	97	10	US-09-864-761-32467	Sequence 32467, A
	34	11.8	59.0	31	10	US-09-801-274-353	Sequence 353, App
c	35	11.8	59.0	31	10	US-09-801-274-958	Sequence 958, App
c	36	11.8	59.0	31	10	US-09-091-134-1	Sequence 1, Appli
	37	11.8	59.0	35	10	US-09-006-298-12	Sequence 12, Appl
c	38	11.8	59.0	78	10	US-09-985-357A-11	Sequence 11, Appl
	39	11.8	59.0	85	10	US-09-919-580-702	Sequence 702, App
c	40	11.8	59.0	88	10	US-09-815-343-316	Sequence 316, App
	41	11.8	59.0	88	10	US-09-919-580-574	Sequence 574, App
c	42	11.8	59.0	93	10	US-09-764-887-545	Sequence 545, App
	43	11.6	58.0	20	10	US-09-791-406-71	Sequence 71, Appl
c	44	11.6	58.0	21	10	US-09-837-235-11	Sequence 11, Appl
c	45	11.6	58.0	21	10	US-09-837-235-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-887-469-13
; Sequence 13, Application US/09887469
; Patent No. US20020146433A1
; GENERAL INFORMATION:
; APPLICANT: Krempl, Christine D.
; APPLICANT: Collins, Peter L.
; APPLICANT: Murphy, Brian R.
; APPLICANT: Buchholz, Ursula
; APPLICANT: Whitehead, Stephen S.
; TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS VACCINES EXPRESSING
; TITLE OF INVENTION: PROTECTIVE ANTIGENS FROM PROMOTOR-PROXIMAL-GENES
; FILE REFERENCE: 15280-424-1US
; CURRENT APPLICATION NUMBER: US/09/887,469
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,708
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 55
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Artificial
; OTHER INFORMATION: Respiratory Syncytial Virus
US-09-887-469-13

Query Match 72.0%; Score 14.4; DB 10; Length 55;
Best Local Similarity 93.8%; Pred. No. 7.8e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3 GCGGCAGCCCCCTCTC 18
           ||||| ||||||||||
Db      22 GCGGCCGCCCCCTCTC 37
```

RESULT 2
US-09-801-274-1010
; Sequence 1010, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:

```
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1010
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-1010

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 31;
  Best Local Similarity   82.4%;  Pred. No. 2e+03;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CGGCAGCCCCCTCTCCA 20
          | |||||||||:|| |||
Db      6 CAGCAGCCCCMTCGCCA 22


RESULT 3
US-09-918-063-35/c
; Sequence 35, Application US/09918063
; Patent No. US20020098546A1
; GENERAL INFORMATION:
;  APPLICANT: Weber, Eric  R.
;  APPLICANT:  Jensen, Wayne A.
;  APPLICANT:  Chandrashekar, Ramaswamy
;  TITLE OF INVENTION: CANINE TAg1 PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
;  FILE REFERENCE: DI-12
;  CURRENT APPLICATION NUMBER: US/09/918,063
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: 60/224,655
;  PRIOR FILING DATE: 2000-08-11
;  NUMBER OF SEQ ID NOS: 85
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
;   LENGTH: 69
;   TYPE: DNA
;   ORGANISM: Canis familiaris
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(69)
;   OTHER INFORMATION:
US-09-918-063-35

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 69;
  Best Local Similarity   83.3%;  Pred. No. 2.4e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTC 18
          ||||||| ||||| || |
Db     56 GAGCGGCGGCCCCTTCGC 39


RESULT 4
US-09-918-063-37
; Sequence 37, Application US/09918063
; Patent No. US20020098546A1
; GENERAL INFORMATION:
;  APPLICANT: Weber, Eric  R.
;  APPLICANT:  Jensen, Wayne A.
;  APPLICANT:  Chandrashekar, Ramaswamy
;  TITLE OF INVENTION: CANINE TAg1 PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
;  FILE REFERENCE: DI-12
;  CURRENT APPLICATION NUMBER: US/09/918,063
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: 60/224,655
;  PRIOR FILING DATE: 2000-08-11
;  NUMBER OF SEQ ID NOS: 85
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
;   LENGTH: 69
;   TYPE: DNA
;   ORGANISM: Canis familiaris
US-09-918-063-37

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 69;
  Best Local Similarity   83.3%;  Pred. No. 2.4e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTC 18
          ||||||| ||||| || |
Db     14 GAGCGGCGGCCCCTTCGC 31


RESULT 5
US-09-783-708-7/c
; Sequence 7, Application US/09783708
; Patent No. US20020058041A1
; GENERAL INFORMATION:
;  APPLICANT: The Regents of the University of California
;  TITLE OF INVENTION: KIDNEY-SPECIFIC TUMOR VACCINE DIRECTED AGAINST KIDNEY TUMOR AN
;  FILE REFERENCE: 306T-897420US
;  CURRENT APPLICATION NUMBER: US/09/783,708
;  CURRENT FILING DATE:  2001-02-13
;  PRIOR APPLICATION NUMBER: 60/182,429
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: 60/182,636
;  PRIOR FILING DATE: 2000-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: primer
US-09-783-708-7

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 27;
  Best Local Similarity   87.5%;  Pred. No. 3.6e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTC 16
          ||||||| || |||||
Db     16 GAGCGGCCGCTCCCTC 1


RESULT 6
US-09-359-672-3/c
; Sequence 3, Application US/09359672
; Patent No. US20020146689A1
; GENERAL INFORMATION:
;  APPLICANT:  BLACKBURN, CATHERINE C
;  APPLICANT:  CHAMBERS, Ian P
;  APPLICANT:  MEDVINSKY, Alexander L
;  APPLICANT:  NIWA, Hitoshi
;  APPLICANT:  SMITH, Austin G
;  TITLE OF INVENTION:  DNA EXPRESSION IN TRANSFECTED CELLS AND ASSAYS CARRIED
;  TITLE OF INVENTION:  OUT IN TRANSFECTED CELLS
;  FILE REFERENCE:  06999.0005-01000
;  CURRENT APPLICATION NUMBER: US/09/359,672
;  CURRENT FILING DATE:  1999-07-23
;  PRIOR APPLICATION NUMBER:  GB 9701492.2
;  PRIOR FILING DATE:  1997-01-24
;  PRIOR APPLICATION NUMBER:  PCT/GB98/00216
;  PRIOR FILING DATE:  1998-01-26
```

```
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 61
;   TYPE: DNA
;   ORGANISM: synthetic construct
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (22)..(60)
US-09-359-672-3

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 61;
  Best Local Similarity   87.5%;  Pred. No. 3.6e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCCTC 16
           | ||||||| |||||
Db      59 GCGCGGCAGCTCCCTC 44


RESULT 7
US-09-861-893-41
; Sequence 41, Application US/09861893
; Patent No. US20020045257A1
; GENERAL INFORMATION:
;  APPLICANT: Feinberg, Andrew
;  APPLICANT:  Strichman-Almashanu, Liora
;  APPLICANT:  Jiang, Shan
;  TITLE OF INVENTION: METHODS FOR ASSAYING GENE IMPRINTING AND
;  TITLE OF INVENTION:  METHYLATED CpG ISLANDS
;  FILE REFERENCE: 01107.00128
;  CURRENT APPLICATION NUMBER: US/09/861,893
;  CURRENT FILING DATE:  2001-05-22
;  PRIOR APPLICATION NUMBER: 60/206,158
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/206,161
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 77
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 41
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-861-893-41

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 100;
  Best Local Similarity   87.5%;  Pred. No. 3.6e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCCTC 16
           | | ||||||||||||
Db      71 GCGGGGCAGCCCCCTC 86


RESULT 8
US-09-895-141-40/c
; Sequence 40, Application US/09895141
; Patent No. US20020120008A1
; GENERAL INFORMATION:
;  APPLICANT: Benzer, Seymour
;  APPLICANT:  Min, Kyung-Tai
;  TITLE OF INVENTION: LIFE EXTENSION OF DROSOPHILA BY A DRUG TREATMENT
;  FILE REFERENCE: 30431.3US01
;  CURRENT APPLICATION NUMBER: US/09/895,141
;  CURRENT FILING DATE:  2001-06-29
;  PRIOR APPLICATION NUMBER: 60/215401
;  PRIOR FILING DATE: 2000-06-29
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: 3'RT-PCR
;   OTHER INFORMATION: Primer Sequence for Hexokinase
US-09-895-141-40

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 27;
  Best Local Similarity   78.9%;  Pred. No. 4.4e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 AGCGGCAGCCCCCTCTCCA 20
           |||  |||| |||| ||||
Db      27 AGCTCCAGCTCCCTATCCA 9


RESULT 9
US-09-908-660-5
; Sequence 5, Application US/09908660
; Patent No. US20020081709A1
; GENERAL INFORMATION:
;  APPLICANT: Bogedain, Christoph
;  APPLICANT:  Maass, Gerd
;  APPLICANT:  Hallek, Michael
;  TITLE OF INVENTION: SYSTEM FOR THE PRODUCTION OF AAV VECTORS
;  FILE REFERENCE: 8484-0057-999
;  CURRENT APPLICATION NUMBER: US/09/908,660
;  CURRENT FILING DATE:  2001-07-18
;  PRIOR APPLICATION NUMBER: 09/214,151
;  PRIOR FILING DATE: 1999-08-06
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 41
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer
US-09-908-660-5

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 41;
  Best Local Similarity   78.9%;  Pred. No. 4.4e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCCTCTCC 19
           ||||||| |||| |  |||
Db       7 GAGCGGCCGCCCGCGGTCC 25


RESULT 10
US-09-801-274-1263/c
; Sequence 1263, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1263
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-1263

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 31;
  Best Local Similarity   81.2%;  Pred. No. 5.3e+03;
```

```
  Matches   13;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GGCAGCCCCCTCTCCA 20
           ||| || |||||||:|
Db      30 GGCCGCACCCTCTCYA 15


RESULT 11
US-09-920-118-1/c
; Sequence 1, Application US/09920118
; Patent No. US20020102734A1
; GENERAL INFORMATION:
;  APPLICANT: Menzel, Rolf
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIRECTED GENE ASSEMBLY
;  FILE REFERENCE: 10424-003
;  CURRENT APPLICATION NUMBER: US/09/920,118
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: 60/222,134
;  PRIOR FILING DATE: 2000-07-31
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 87
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence: Primer
US-09-920-118-1

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 87;
  Best Local Similarity   92.9%;  Pred. No. 5.3e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCC 14
           ||||||| ||||||
Db      72 GAGCGGCCGCCCCC 59


RESULT 12
US-09-920-118-2
; Sequence 2, Application US/09920118
; Patent No. US20020102734A1
; GENERAL INFORMATION:
;  APPLICANT: Menzel, Rolf
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIRECTED GENE ASSEMBLY
;  FILE REFERENCE: 10424-003
;  CURRENT APPLICATION NUMBER: US/09/920,118
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: 60/222,134
;  PRIOR FILING DATE: 2000-07-31
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 87
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence: Primer
US-09-920-118-2

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 87;
  Best Local Similarity   92.9%;  Pred. No. 5.3e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCC 14
           ||||||| ||||||
Db      20 GAGCGGCCGCCCCC 33


RESULT 13
US-10-037-616-17
; Sequence 17, Application US/10037616
; Patent No. US20020123148A1
; GENERAL INFORMATION:
;  APPLICANT: English, Denis
;  APPLICANT:  Kovacs, Richard J.
;  APPLICANT:  Rizzo, Maria T.
;  APPLICANT:  Sliva, Daniel T.
;  TITLE OF INVENTION: Sphingolipid Compositions and Methods for Their Therapeutic Us
;  FILE REFERENCE: 7042-119
;  CURRENT APPLICATION NUMBER: US/10/037,616
;  CURRENT FILING DATE:  2002-04-19
;  PRIOR APPLICATION NUMBER: US 60/243,887
;  PRIOR FILING DATE: 2000-10-27
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
;   LENGTH: 18
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense oligonucleotide
US-10-037-616-17

  Query Match             61.0%;  Score 12.2;  DB 12;  Length 18;
  Best Local Similarity   82.4%;  Pred. No. 6.4e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 CGGCAGCCCCCTCTCCA 20
           | |||||||||  |||||
Db       1 CAGCAGCCCCGACTCCA 17


RESULT 14
US-09-732-348-8/c
; Sequence 8, Application US/09732348
; Patent No. US20020046419A1
; GENERAL INFORMATION:
;  APPLICANT:  Yen Choo, et al.
;  TITLE OF INVENTION:   Regulated Gene Expression in Plants
;  FILE REFERENCE:   674538-2001
;  CURRENT APPLICATION NUMBER: US/09/732,348
;  CURRENT FILING DATE:  2000-12-07
;  NUMBER OF SEQ ID NOS:   21
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 8
;   LENGTH: 35
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(35)
;   OTHER INFORMATION: forward primer
US-09-732-348-8

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 35;
  Best Local Similarity   82.4%;  Pred. No. 6.4e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 AGCGGCAGCCCCCTCTC 18
           | ||||||| || ||||
Db      33 ACCGGCAGCGCCTTCTC 17


RESULT 15
US-09-732-348-20/c
; Sequence 20, Application US/09732348
; Patent No. US20020046419A1
; GENERAL INFORMATION:
;  APPLICANT:  Yen Choo, et al.
;  TITLE OF INVENTION:   Regulated Gene Expression in Plants
;  FILE REFERENCE:   674538-2001
;  CURRENT APPLICATION NUMBER: US/09/732,348
;  CURRENT FILING DATE:  2000-12-07
;  NUMBER OF SEQ ID NOS:   21
;  SOFTWARE:   PatentIn version 3.0
```

; SEQ ID NO 20
; LENGTH: 38
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(38)
; OTHER INFORMATION: forward primer
US-09-732-348-20

Query Match 61.0%; Score 12.2; DB 10; Length 38;
Best Local Similarity 82.4%; Pred. No. 6.4e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 AGCGGCAGCCCCCTCTC 18
           | ||||||| || ||||
Db      36 ACCGGCAGCGCCTTCTC 20
```

Search completed: November 23, 2002, 06:42:00
Job time : 18.25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-1
Perfect score: 20
Sequence: 1 gagcggcagccccctctcca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	15.2	76.0	98	10	BE425297	BE425297 WHE0311_D
	2	14.8	74.0	65	9	AI270452	AI270452 qu88c06.x
	3	14.2	71.0	50	9	AU107434	AU107434 AU107434
	4	14.2	71.0	58	9	AI123506	AI123506 qa01g09.x
c	5	14.2	71.0	59	17	BH855612	BH855612 SALK_0848
	6	14.2	71.0	85	14	BQ808288	BQ808288 1030003A0
c	7	14.2	71.0	87	17	TA2H03P	AL441676 T. brucei
	8	13.8	69.0	74	12	BG315169	BG315169 PO1.0.136
	9	13.8	69.0	84	13	BM090905	BM090905 ig17d06.x
	10	13.8	69.0	86	10	AW593901	AW593901 hg28h03.x
	11	13.8	69.0	95	9	AI051259	AI051259 oy49f02.x
	12	13.8	69.0	96	9	AA835222	AA835222 ak65b02.s
c	13	13.8	69.0	97	17	AZ356019	AZ356019 1M0095M22
	14	13.8	69.0	100	10	AW605663	AW605663 QV2-DT007
c	15	13.6	68.0	57	17	AZ962155	AZ962155 2M0230D24
	16	13.6	68.0	66	17	TA391A11P	AL498916 T. brucei
c	17	13.6	68.0	67	13	BI908580	BI908580 603069979
c	18	13.6	68.0	82	9	AA870618	AA870618 vq24a01.r
c	19	13.6	68.0	89	9	AA579142	AA579142 nf35g06.s
c	20	13.4	67.0	75	17	AZ339132	AZ339132 1M0070B03
c	21	13.4	67.0	93	17	TA358C08Q	AL495379 T. brucei
c	22	13.2	66.0	43	17	BH855596	BH855596 SALK_0849
	23	13.2	66.0	50	9	AU103490	AU103490 AU103490
c	24	13.2	66.0	61	17	AZ791343	AZ791343 2M0041A16
	25	13.2	66.0	79	9	AI799649	AI799649 to74d11.x
	26	13.2	66.0	79	12	BE880374	BE880374 601490852
	27	13.2	66.0	82	9	AI915454	AI915454 tr28d07.x
c	28	13.2	66.0	82	12	BF168237	BF168237 601773908
	29	13.2	66.0	90	9	AA680840	AA680840 LmFrAm039
c	30	13.2	66.0	93	10	BE287079	BE287079 601091066
c	31	13	65.0	60	17	BH891314	BH891314 3526_1_17
c	32	12.8	64.0	37	9	AA897070	AA897070 al06h07.s
	33	12.8	64.0	50	9	AU102552	AU102552 AU102552
	34	12.8	64.0	50	9	AU102553	AU102553 AU102553
	35	12.8	64.0	50	9	AU102554	AU102554 AU102554
c	36	12.8	64.0	53	12	BF013395	BF013395 rk31b05.y
	37	12.8	64.0	61	12	BE740294	BE740294 601595279
c	38	12.8	64.0	70	13	BJ084169	BJ084169 BJ084169
	39	12.8	64.0	86	17	BH413188	BH413188 1007031B0
c	40	12.8	64.0	87	14	BQ821696	BQ821696 1030094C0
	41	12.8	64.0	87	17	FR0023597	AL016455 F.rubripe
c	42	12.8	64.0	91	9	AA865904	AA865904 oh50g06.s
	43	12.8	64.0	91	12	BF436210	BF436210 nab45d10.
c	44	12.8	64.0	93	12	BF506824	BF506824 10J22L59b
c	45	12.8	64.0	94	14	BQ808815	BQ808815 1030007C0

ALIGNMENTS

RESULT 1
BE425297/c
LOCUS BE425297 98 bp mRNA linear EST 24-JUL-2000
DEFINITION WHE0311_D02_D02ZS Wheat unstressed seedling shoot cDNA library Triticum aestivum cDNA clone WHE0311_D02_D02, mRNA sequence.
ACCESSION BE425297
VERSION BE425297.1 GI:9423056
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 98)
AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.
TITLE The structure and function of the expressed portion of the wheat genomes
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20

```
          Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE0311_D02_D02"
                     /clone_lib="Wheat unstressed seedling shoot cDNA library"
                     /tissue_type="Etiolated shoot"
                     /dev_stage="Five day old seedling"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized
                     , germinated and grown aseptically in the  dark at room
                     temperature on filter paper with water, nystatin and
                     cefotaxime in covered crystallization dishes. Shoots were
                     harvested.  The tissue, total RNA, and poly(A) RNA were
                     prepared, a cDNA library  was made, and the cDNA clones
                     were in vivo excised to give pBluescript  phagemids in the
                     TJ Close lab (Choi, Close, Fenton) at the University  of
                     California, Riverside. Plasmid DNA preparations and DNA
                     sequencing  were performed in the OD Anderson lab (all
                     other authors)."
BASE COUNT       17 a     30 c     36 g     15 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 10;  Length 98;
  Best Local Similarity   85.0%;  Pred. No. 2.3e+04;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCCTCTCCA 20
           |||||||||||||  | |||
Db      94 GAGCGGCAGCCCCTGCACCA 75


RESULT 2
AI270452
LOCUS       AI270452                 65 bp    mRNA    linear   EST 17-NOV-1998
DEFINITION  qu88c06.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:1979146 3',
            mRNA sequence.
ACCESSION   AI270452
VERSION     AI270452.1  GI:3889619
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1979146"
                     /clone_lib="NCI_CGAP_Gas4"
                     /tissue_type="poorly differentiated adenocarcinoma with
                     signet ring cell features"
                     /lab_host="DH10B"
                     /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.69 kb. Life Technologies catalog #:
                     11549-011"
BASE COUNT       10 a     27 c     18 g     10 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 9;  Length 65;
  Best Local Similarity   88.9%;  Pred. No. 3.2e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 AGCGGCAGCCCCCTCTCC 19
           |||||| |||||||| ||
Db       1 AGCGGCCGCCCCCTCGCC 18


RESULT 3
AU107434
LOCUS       AU107434                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU107434 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            LNG15774, mRNA sequence.
ACCESSION   AU107434
VERSION     AU107434.1  GI:13556955
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="LNG15774"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT       10 a     23 c     14 g      3 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 50;
  Best Local Similarity   84.2%;  Pred. No. 5.5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCCTCTCC 19
           |||||| ||| || |||||
Db      13 GAGCGGGAGCGCCTTCTCC 31


RESULT 4
AI123506
LOCUS       AI123506                 58 bp    mRNA    linear   EST 01-OCT-1998
DEFINITION  qa01g09.x1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
            IMAGE:1683616 3' similar to SW:DIA_DROME P48608 DIAPHANOUS PROTEIN.
            ;contains element MSR1 repetitive element ;, mRNA sequence.
ACCESSION   AI123506
VERSION     AI123506.1  GI:3539272
KEYWORDS    EST.
SOURCE      human.
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
            , Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 953   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1683616"
                     /clone_lib="Soares_parathyroid_tumor_NbHPA"
                     /tissue_type="parathyroid tumor"
                     /dev_stage="adult"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTTT
                     TTTTT-3'], double-stranded cDNA was size selected, ligated
                     to Eco RI adapters (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of a modified pT7T3
                     vector (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
                     adenomas was kindly provided by Dr. Stephen Marx, National
                     Institute of Diabetes and Digestive and Kidney Diseases,
                     NIH."
BASE COUNT       18 a      24 c      13 g       3 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 58;
  Best Local Similarity   84.2%;  Pred. No. 5.5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 AGCGGCAGCCCCCTCTCCA 20
           ||| |||||||||| | |||
Db      10 AGCCGCAGCCCCCCCCCCA 28


RESULT 5
BH855612/c
LOCUS       BH855612                 59 bp    DNA     linear   GSS 08-JUL-2002
DEFINITION  SALK_084890.49.05.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_084890.49.05.x, DNA
            sequence.
ACCESSION   BH855612
VERSION     BH855612.1  GI:21705202
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated exon of At4g31410.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_084890.49.05.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       21 a      7 c      20 g      11 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 59;
  Best Local Similarity   84.2%;  Pred. No. 5.5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GAGCGGCAGCCCCCTCTCC 19
           |||||||  | ||||||||
Db      45 GAGCGGCGACTCCCTCTCC 27


RESULT 6
BQ808288
LOCUS       BQ808288                 85 bp    mRNA    linear   EST 01-AUG-2002
DEFINITION  1030003A09.x2 C. reinhardtii CC-1690, Deflagellation (normalized),
            Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BQ808288
VERSION     BQ808288.1  GI:22048620
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
            ,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1030
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Chlamydomonas reinhardtii"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, Deflagellation
                     (normalized), Lambda Zap II"
```

/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; Deflagellation library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells which had been re-synthesizing flagella for 15, 30 and 60 min after being deflagellated by pH shock. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT 16 a 39 c 30 g 0 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 85;
Best Local Similarity 84.2%; Pred. No. 5.7e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 GAGCGGCAGCCCCCTCTCC 19
          ||||||||| |||| | ||
Db     35 GAGCGGCAGACCCCCCCCC 53
```

RESULT 7
TA2H03P/c
LOCUS TA2H03P 87 bp DNA linear GSS 16-SEP-2000
DEFINITION T. brucei sheared genomic DNA clone 2h03, forward sequence, genomic survey sequence.
ACCESSION AL441676
VERSION AL441676.1 GI:10185383
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 87)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nhl@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v+i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES Location/Qualifiers
source 1. .87
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="2h03"
BASE COUNT 17 a 12 c 41 g 17 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 17; Length 87;
Best Local Similarity 84.2%; Pred. No. 5.7e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      2 AGCGGCAGCCCCCTCTCCA 20
          | ||||| || ||||||||      .
Db     53 ATCGGCATCCTCCTCTCCA 35
```

RESULT 8
BG315169
LOCUS BG315169 74 bp mRNA linear EST 28-FEB-2002
DEFINITION PO1.0.136 Human THP1 cell line library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG315169
VERSION BG315169.1 GI:18998041
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 74)
AUTHORS Andersson,T., Borang,S., Larsson,M., Thelin,A., Ekstrand-Hammarsrom,B., Wirta,V., Wenborg,A., Lundeberg,J. and Odeberg,J.
TITLE Identification of candidate genes in atherosclerosis - Virtual chip analysis in RDA based transcript profiling of monocyte/macrophage response to oxidised LDL
JOURNAL Unpublished (2001)
COMMENT Contact: Andersson Tove
Department of Biotechnology
KTH
Teknikringen 34, plan 6, 100 44 Stockholm, Sweden
Tel: +46 8 790 71 29
Fax: +46 8 245452
Email: tove@biochem.kth.se
POLYA=No.
FEATURES Location/Qualifiers
source 1. .74
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human THP1 cell line library"
/cell_line="THP1"
/note="Vector: pRIT28; Site_1: BamHI; Site_2: BamHI; Shotgun cloning of RDA difference products. Macrophage and foamcell libraries were submitted to successive rounds of subtractive hybridisations generating populations of gene fragments that are differentially expressed in macrophage to foam cell formation."
BASE COUNT 15 a 25 c 21 g 12 t 1 others
ORIGIN

Query Match 69.0%; Score 13.8; DB 12; Length 74;
Best Local Similarity 88.2%; Pred. No. 8.1e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      2 AGCGGCAGCCCCCTCTC 18
          ||||||||| |||| |
Db     21 AGCGGCAGCCACCTCCC 37
```

RESULT 9
BM090905
LOCUS BM090905 84 bp mRNA linear EST 20-NOV-2001
DEFINITION ig17d06.x1 Human Fetal Pancreas 1A Homo sapiens cDNA 3' similar to TR:Q29269 Q29269 UNKNOWN PROTEIN ;, mRNA sequence.
ACCESSION BM090905
VERSION BM090905.1 GI:17019871
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 84)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium

```
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: ig17d06.y1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: Dr. Hiroshi Inoue
            (hinoue@imgate.wustl.edu).
FEATURES              Location/Qualifiers
     source           1. .84
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="Human Fetal Pancreas 1A"
                      /tissue_type="Fetal Pancreas (4 Pooled Donors, 18 - 20
                      weeks, Stratagene #738023)"
                      /dev_stage="Fetal Pancreas"
                      /note="Vector: pBluescript SK(-); Site_1: NotI; Site_2:
                      XhoI; cDNA made by oligo-dT priming. Size-selected by
                      CHROMA SPIN-1000 (Clontech). Average insert size ~1kb. 5'
                      XhoI site was destroyed after directional cloning.
                      Amplified once. Contact information: Hiroshi Inoue, MD,
                      Metabolism Div. (Alan Permutt Lab), Washington University
                      School of Medicine, Box 8127, 660 South Euclid Ave., St.
                      Louis, MO 63110 E-mail: hinoue@imgate.wustl.edu Tel:
                      314-362-1916, Fax:314-747-2692."
BASE COUNT       12 a     32 c     25 g     15 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 13;  Length 84;
  Best Local Similarity   88.2%;  Pred. No. 8.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 GCGGCAGCCCCCTCTCC 19
          ||||| |||||||| ||
Db      2 GCGGCCGCCCCCTCGCC 18


RESULT 10
AW593901
LOCUS       AW593901                 86 bp    mRNA    linear   EST 22-MAR-2000
DEFINITION  hg28h03.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2946965 3',
            mRNA sequence.
ACCESSION   AW593901
VERSION     AW593901.1  GI:7281159
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml

            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .86
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2946965"
                      /clone_lib="NCI_CGAP_GC6"
                      /tissue_type="pooled germ cell tumors"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
                      from the normalized library NCI_CGAP_GC4 was prepared, and
                      ss circles were made in vitro. Following HAP purification,
                      this DNA was used as tracer in a subtractive hybridization
                      reaction. The driver was PCR-amplified cDNAs from a pool
                      of 5,000 clones made from the same library (cloneIDs
                      1257096-1258631, 1469064-1470983, and 1475592-1476743).
                      Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       12 a     32 c     25 g     17 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 86;
  Best Local Similarity   88.2%;  Pred. No. 8.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 GCGGCAGCCCCCTCTCC 19
          ||||| |||||||| ||
Db      1 GCGGCCGCCCCCTCGCC 17


RESULT 11
AI051259
LOCUS       AI051259                 95 bp    mRNA    linear   EST 24-SEP-1998
DEFINITION  oy49f02.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:1669179 3',
            mRNA sequence.
ACCESSION   AI051259
VERSION     AI051259.1  GI:3306793
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 843   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES              Location/Qualifiers
     source           1. .95
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1669179"
                      /clone_lib="NCI_CGAP_Brn23"
                      /tissue_type="glioblastoma (pooled)"
                      /lab_host="DH10B"
                      /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTTTT
```

T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 16 a 35 c 26 g 18 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 95;
Best Local Similarity 88.2%; Pred. No. 8.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GCGGCAGCCCCCTCTCC 19
||||| |||||||| ||
Db 1 GCGGCCGCCCCCTCGCC 17

RESULT 12
AA835222
LOCUS AA835222 96 bp mRNA linear EST 23-FEB-1998
DEFINITION ak65b02.s1 Barstead pancreas HPLRB1 Homo sapiens cDNA clone IMAGE:1412715 3', mRNA sequence.
ACCESSION AA835222
VERSION AA835222.1 GI:2908950
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 96)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: polyT not found
Seq primer: -40m13 fwd. ET from Amersham.
FEATURES Location/Qualifiers
source 1. .96
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1412715"
/clone_lib="Barstead pancreas HPLRB1"
/sex="female"
/dev_stage="adult, 34 years"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGGATCCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead. "
BASE COUNT 13 a 35 c 29 g 19 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 96;
Best Local Similarity 88.2%; Pred. No. 8.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GCGGCAGCCCCCTCTCC 19
||||| |||||||| ||
Db 1 GCGGCCGCCCCCTCGCC 17

RESULT 13
AZ356019/c
LOCUS AZ356019 97 bp DNA linear GSS 02-OCT-2000
DEFINITION 1M0095M22R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0095M22 R, DNA sequence.
ACCESSION AZ356019
VERSION AZ356019.1 GI:10468908
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 97)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0095 row: M column: 22
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 97.
FEATURES Location/Qualifiers
source 1. .97
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0095M22"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 15 a 23 c 46 g 13 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 97;
Best Local Similarity 88.2%; Pred. No. 8.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GCGGCAGCCCCCTCTCC 19
||||||| |||||| ||

```
Db     59 GCGGCAGTCCCCTCGCC 43


RESULT 14
AW605663
LOCUS       AW605663                100 bp    mRNA    linear   EST 23-MAR-2000
DEFINITION  QV2-DT0075-190100-027-a12 DT0075 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW605663
VERSION     AW605663.1  GI:7310404
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV2&t2=QV2-DT0075-
            190100-027-a12&t3=2000-01-19&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 100.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="DT0075"
                     /dev_stage="Adult"
                     /note="Organ: denis_drash; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       16 a     35 c     32 g     17 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 100;
  Best Local Similarity   88.2%;  Pred. No. 8.3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 GCGGCAGCCCCCTCTCC 19
          ||||| |||||||| ||
Db     40 GCGGCCGCCCCCTCGCC 56


RESULT 15
AZ962155/c
LOCUS       AZ962155                  57 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0230D24R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0230D24 R, DNA sequence.
ACCESSION   AZ962155
VERSION     AZ962155.1  GI:13833382
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0230  row: D  column: 24
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 57.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0230D24"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        9 a     10 c     23 g     15 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 57;
  Best Local Similarity   80.0%;  Pred. No. 9.5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GAGCGGCAGCCCCCTCTCCA 20
          ||  |||||| || ||||||
Db     37 GAAGGGCAGCACCATCTCCA 18


Search completed: November 26, 2002, 04:06:14
Job time : 774.8 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-2
Perfect score: 20
Sequence: 1 cgagcacggcgcagaggagc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		15.4	77.0	99	6	AX435605	AX435605 Sequence
2		15.2	76.0	27	6	AX356224	AX356224 Sequence
3	c	15.2	76.0	87	6	E27048	E27048 Novel recep
4		14.4	72.0	21	6	A91560	A91560 Sequence 87
5	c	14.4	72.0	21	6	AX454812	AX454812 Sequence
6		14.2	71.0	50	6	AX156996	AX156996 Sequence
7		14.2	71.0	51	6	AX156995	AX156995 Sequence
8	c	14.2	71.0	100	10	RATMAP	M26758 Rat major a
9	c	13.8	69.0	37	6	AR170305	AR170305 Sequence
10	c	13.8	69.0	37	6	AX007133	AX007133 Sequence
11	c	13.8	69.0	37	6	E51079	E51079 Novel Esche
12		13.6	68.0	23	6	AR112201	AR112201 Sequence
13		13.6	68.0	23	6	AR149243	AR149243 Sequence
14		13.6	68.0	27	6	I27230	I27230 Sequence 36
15	c	13.6	68.0	35	6	A07810	A07810 Oligonucleo
16		13.6	68.0	43	6	A07815	A07815 Oligonucleo
17		13.6	68.0	45	6	AR012045	AR012045 Sequence
18		13.6	68.0	45	6	AR062287	AR062287 Sequence
19		13.6	68.0	50	6	AX014759	AX014759 Sequence
20	c	13.6	68.0	60	6	I25129	I25129 Sequence 17
21	c	13.6	68.0	75	6	AR078717	AR078717 Sequence
22	c	13.6	68.0	75	6	AR198697	AR198697 Sequence
23	c	13.6	68.0	75	6	I25146	I25146 Sequence 42
24		13.6	68.0	75	8	MZEMTATP92	L38813 Maize mitoc
25	c	13.6	68.0	81	6	A10907	A10907 DNA for sec
26	c	13.6	68.0	81	6	AX014760	AX014760 Sequence
27	c	13.6	68.0	98	9	HUMTFAA1	M17610 Human trans
28	c	13.2	66.0	41	6	AX327061	AX327061 Sequence
29	c	13.2	66.0	44	6	A85333	A85333 Sequence 13
30		13.2	66.0	51	6	AX158837	AX158837 Sequence
31		13.2	66.0	92	6	I72525	I72525 Sequence 10
32	c	13.2	66.0	99	8	APA228812	AJ228812 Alstroeme
33		13	65.0	30	6	E38400	E38400 NF-kappa B
34		13	65.0	37	6	I43910	I43910 Sequence 12
35	c	13	65.0	45	6	I43911	I43911 Sequence 13
36		12.8	64.0	20	6	AR124444	AR124444 Sequence
37	c	12.8	64.0	24	6	AX290848	AX290848 Sequence
38		12.8	64.0	75	6	AR092891	AR092891 Sequence
39		12.8	64.0	75	6	AX133309	AX133309 Sequence
40	c	12.6	63.0	30	6	AR135861	AR135861 Sequence
41	c	12.6	63.0	33	6	AR134695	AR134695 Sequence
42	c	12.6	63.0	42	6	E04535	E04535 Oligonucleo
43		12.6	63.0	44	17	AF294387	Af294387 Homo sapi
44		12.6	63.0	50	6	AX161340	AX161340 Sequence
45	c	12.6	63.0	51	6	AX159219	AX159219 Sequence

ALIGNMENTS

RESULT 1
AX435605
LOCUS AX435605 99 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 4020 from Patent WO0229113.
ACCESSION AX435605
VERSION AX435605.1 GI:21660413
KEYWORDS .
SOURCE Bacillus licheniformis.
ORGANISM Bacillus licheniformis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Berka,R. and Clausen,I.G.
TITLE Methods for monitoring multiple gene expression
JOURNAL Patent: WO 0229113-A 4020 11-APR-2002;
Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)

```
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Bacillus licheniformis"
                     /db_xref="taxon:1402"
BASE COUNT       32 a      26 c      20 g      20 t       1 others
ORIGIN

  Query Match             77.0%;  Score 15.4;  DB 6;  Length 99;
  Best Local Similarity   94.1%;  Pred. No. 2.6e+04;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GCACGGCGCAGAGGAGC 20
           ||||| ||||||||||
Db      31 GCACGCCGCAGAGGAGC 47


RESULT 2
AX356224
LOCUS       AX356224                  27 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 18 from Patent WO0200905.
ACCESSION   AX356224
VERSION     AX356224.1  GI:18620731
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Conner,T.W., Dubois,P., Malven,M. and Masucci,J.D.
  TITLE     Plant regulatory sequences for selective control of gene expression
  JOURNAL   Patent: WO 0200905-A 18 03-JAN-2002;
            Monsanto Technology LLC (US)
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
BASE COUNT        8 a       6 c      11 g       2 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 6;  Length 27;
  Best Local Similarity   85.0%;  Pred. No. 3.8e+04;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           |||| |||||| |||||| |
Db       7 CGAGGACGGCGAAGAGGATC 26


RESULT 3
E27048/c
LOCUS       E27048                    87 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Novel receptor protein and utilization of the same.
ACCESSION   E27048
VERSION     E27048.1  GI:13026378
KEYWORDS    JP 1999152300-A/3.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   Kazunori,N., Yasushi,A. and Takashi,H.
  TITLE     Novel receptor protein and utilization of the same
  JOURNAL   Patent: JP 1999152300-A 3 08-JUN-1999;
            TAKEDA CHEM IND LTD
COMMENT     OS   Unidentified
            PN   JP 1999152300-A/3
            PD   08-JUN-1999
            PF   24-APR-1998 JP 1998114450
            PR
            PI   KAZUNORI NISHI,YASUSHI ARAYA,TAKASHI HORIGUCHI PC
            C07K14/715,A61K45/00,C07K16/28,C12N15/09,C12P21/02, PC
            C12Q1/02//A61K48/00,
            PC   C12N5/10,(C12N15/09,C12R1:91),(C12P21/02,C12R1:91),C12N15/00,
            PC   C12N5/00,
            PC   (C12N15/00,C12R1:91)
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .87
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        3 a      36 c      32 g      16 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 6;  Length 87;
  Best Local Similarity   85.0%;  Pred. No. 3.2e+04;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           |||||| |||||||| |||
Db      64 CGAGCAGCGCGCAGAGCAGC 45


RESULT 4
A91560
LOCUS       A91560                    21 bp    DNA     linear   PAT 22-JAN-2000
DEFINITION  Sequence 87 from Patent WO9824928.
ACCESSION   A91560
VERSION     A91560.1  GI:6740515
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Pallisgaard,N. and Hokland,P.
  TITLE     DETECTION OF CHROMOSOMAL ABNORMALITIES
  JOURNAL   Patent: WO 9824928-A 87 11-JUN-1998;
            PALLISGAARD NIELS (DK); HOKLAND PETER (DK)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        7 a       3 c       9 g       2 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 6;  Length 21;
  Best Local Similarity   93.8%;  Pred. No. 8.6e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 AGCACGGCGCAGAGGA 18
           ||||||| ||||||||
Db       1 AGCACGGAGCAGAGGA 16


RESULT 5
AX454812/c
LOCUS       AX454812                  21 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 14 from Patent WO0206535.
ACCESSION   AX454812
VERSION     AX454812.1  GI:21714041
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Miller,M.M., Afanassieff,M. and Briles,W.E.
  TITLE     Method for breeding disease resistant domesticated fowl
  JOURNAL   Patent: WO 0206535-A 14 24-JAN-2002;
            City of Hope (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer for SSCP analysis of Mhc B class II-beta
```

```
                    genotype of fowl"
BASE COUNT        4 a      7 c      6 g      4 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 6;  Length 21;
  Best Local Similarity   93.8%;  Pred. No. 8.6e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 CACGGCGCAGAGGAGC 20
          ||||| ||||||||||
Db     16 CACGGTGCAGAGGAGC 1


RESULT 6
AX156996
LOCUS       AX156996                 50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 324 from Patent WO0140521.
ACCESSION   AX156996
VERSION     AX156996.1  GI:14538327
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 324 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg20436198"
     misc_feature    26
                     /note="2 of 2 allelic variants (323 is other entry)"
BASE COUNT        5 a      14 c     28 g      3 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 50;
  Best Local Similarity   84.2%;  Pred. No. 9.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAG 19
          |||| ||||||| || |||
Db     30 CGAGGACGGCGCGGACGAG 48


RESULT 7
AX156995
LOCUS       AX156995                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 323 from Patent WO0140521.
ACCESSION   AX156995
VERSION     AX156995.1  GI:14538326
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 323 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (324 is other entry)
                     Accession number cg20436198"
BASE COUNT        5 a      14 c     29 g      3 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 9.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAG 19
          |||| ||||||| || |||
Db     31 CGAGGACGGCGCGGACGAG 49


RESULT 8
RATMAP/c
LOCUS       RATMAP                   100 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Rat major acute phase protein mRNA (alpha1-MAP gene), partial cds.
ACCESSION   M26758
VERSION     M26758.1  GI:205301
KEYWORDS    major acute phase protein.
SOURCE      Rat (strain Sprague-Dawley) liver cDNA to mRNA.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Anderson,K.P., Martin,A.D. and Heath,E.C.
  TITLE     Rat major acute-phase protein: biosynthesis and characterization of
            cDNA clone
  JOURNAL   Arch. Biochem. Biophys. 233 (2), 624-635 (1984)
  MEDLINE   85021400
   PUBMED   6207775
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             <1. .>100
                     /note="major acute phase protein"
                     /codon_start=2
                     /protein_id="AAA41567.1"
                     /db_xref="GI:205302"
                     /translation="TILLLCSKLLPSLAQEEGAQEMDRNDETVFQAV"
BASE COUNT       26 a      28 c     24 g     22 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 100;
  Best Local Similarity   84.2%;  Pred. No. 8.2e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GAGCACGGCGCAGAGGAGC 20
          ||||  || ||||||||||
Db     28 GAGCTTGGAGCAGAGGAGC 10


RESULT 9
AR170305/c
LOCUS       AR170305                 37 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 19 from patent US 6291245.
ACCESSION   AR170305
VERSION     AR170305.1  GI:17908264
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Kopetzki,E. and Schantz,C.
  TITLE     Host-vector system
  JOURNAL   Patent: US 6291245-A 19 18-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="unknown"
BASE COUNT        8 a      11 c     11 g      7 t
```

```
ORIGIN

  Query Match           69.0%;  Score 13.8;  DB 6;  Length 37;
  Best Local Similarity  88.2%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGG 17
          |||||||||||||  ||
Db     28 CGAGCACGGCGCACCGG 12


RESULT 10
AX007133/c
LOCUS       AX007133                37 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 19 from Patent EP0972838.
ACCESSION   AX007133
VERSION     AX007133.1  GI:9995024
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Schantz,C. and Kopetzki,E.
  TITLE     Escherichia coli host/vector system based on antibiotic-free
            selection by complementation of an auxotrophy
  JOURNAL   Patent: EP 0972838-A 19 19-JAN-2000;
            ROCHE DIAGNOSTICS GMBH (DE)
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer"
BASE COUNT        8 a     11 c     11 g      7 t
ORIGIN

  Query Match           69.0%;  Score 13.8;  DB 6;  Length 37;
  Best Local Similarity  88.2%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGG 17
          |||||||||||||  ||
Db     28 CGAGCACGGCGCACCGG 12


RESULT 11
E51079/c
LOCUS       E51079                  37 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Novel Escherichia coli/host vector system based on selection not
            relying on antibiotic by complementing auxotrophy.
ACCESSION   E51079
VERSION     E51079.1  GI:13023276
KEYWORDS    JP 2000050888-A/19.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Erhard,K. and Chrithtiane,S.
  TITLE     Novel Escherichia coli/host vector system based on selection not
            relying on antibiotic by complementing auxotrophy
  JOURNAL   Patent: JP 2000050888-A 19 22-FEB-2000;
            ROCHE DIAGNOSTICS GMBH
COMMENT     OS   Artificial Sequence
            PN   JP 2000050888-A/19
            PD   22-FEB-2000
            PF   15-JUL-1999 JP 1999202303
            PR   15-JUL-1998 EP   98113156.8,09-OCT-1998 EP   98119078.8 PI
            ERHARD KOPETSUKI,CHRITHTIANE SCHANZ
            PC   C12N15/09,C12N1/21,C12P21/02//(C12N1/21,C12R1:19),(C12P21/02,
            PC   C12R1:19),
            PC   C12N15/00
            CC
            FH   Key             Location/Qualifiers
            FT   source          1. .37
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        8 a     11 c     11 g      7 t
ORIGIN

  Query Match           69.0%;  Score 13.8;  DB 6;  Length 37;
  Best Local Similarity  88.2%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGG 17
          |||||||||||||  ||
Db     28 CGAGCACGGCGCACCGG 12


RESULT 12
AR112201
LOCUS       AR112201                23 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 90 from patent US 6130041.
ACCESSION   AR112201
VERSION     AR112201.1  GI:14092101
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Acton,S.Laurene.
  TITLE     Human intronic and polymorphic SR-BI nucleic acids and uses
            therefor
  JOURNAL   Patent: US 6130041-A 90 10-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="unknown"
BASE COUNT        6 a     10 c      6 g      1 t
ORIGIN

  Query Match           68.0%;  Score 13.6;  DB 6;  Length 23;
  Best Local Similarity  80.0%;  Pred. No. 1.8e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          | |||||  |||||  ||  |||
Db      4 CCAGCACAGCGCACAGTAGC 23


RESULT 13
AR149243
LOCUS       AR149243                23 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 90 from patent US 6228581.
ACCESSION   AR149243
VERSION     AR149243.1  GI:15113834
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Acton,S.L. and Ordovas,J.M.
  TITLE     Human intronic and polymorphic SR-BI nucleic acids and uses
            therefor
  JOURNAL   Patent: US 6228581-A 90 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="unknown"
BASE COUNT        6 a     10 c      6 g      1 t
ORIGIN

  Query Match           68.0%;  Score 13.6;  DB 6;  Length 23;
  Best Local Similarity  80.0%;  Pred. No. 1.8e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
```

```
          | ||||| ||||| || |||
Db      4 CCAGCACAGCGCACAGTAGC 23


RESULT 14
I27230
LOCUS       I27230                    27 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 36 from patent US 5563254.
ACCESSION   I27230
VERSION     I27230.1  GI:1818006
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Hoffman,S.J. and Nagai,K.
  TITLE     Polynucleotides encoding genetically engineered mutant hemoglobins
  JOURNAL   Patent: US 5563254-A 36 08-OCT-1996;
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unknown"
BASE COUNT        7 a      11 c      7 g      2 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 27;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          | ||||||  |||  ||||||
Db      6 CCAGCACGTCGCCCAGGAGC 25


RESULT 15
A07810/c
LOCUS       A07810                    35 bp    DNA     linear   PAT 26-AUG-1993
DEFINITION  Oligonucleotide.
ACCESSION   A07810
VERSION     A07810.1  GI:413246
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Paques,E.P.
  TITLE     Mutants of urinary plasminogen activator, their production and use
  JOURNAL   Patent: EP 0387380-A 1 19-SEP-1990;
            BEHRINGWERKE Aktiengesellschaft
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        2 a      13 c      11 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 35;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          ||| || | ||||||| |||
Db     27 CGACCAGGACGCAGAGAAGC 8


Search completed: December  3, 2002, 18:12:02
Job time : 312.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-2
Perfect score: 20
Sequence: 1 cgagcacggcgcagaggagc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31432	Human neuropilin m
	2	15.4	77.0	99	24	ABK76729	Bacillus lichenifo
	3	15.2	76.0	27	24	ABK27825	Corn male reproduc
	4	14.8	74.0	48	9	AAN80520	Synthetic oligonuc
	5	14.4	72.0	21	19	AAV40985	Primer TELAML1:121
c	6	14.4	72.0	21	24	ABK88264	Exon 2 (alphal dom
c	7	14.4	72.0	21	24	AAD29196	Fowl MHc B class I
	8	14.2	71.0	50	22	AAI73383	Human silent SNP c
	9	14.2	71.0	51	21	AAA77070	Human clone cg4392
	10	14.2	71.0	51	22	AAI73382	Human silent SNP c
c	11	13.8	69.0	37	21	AAZ57918	Escherichia coli t
	12	13.8	69.0	65	24	ABN56998	Mouse spliced tran
	13	13.6	68.0	23	20	AAX24557	Human SR-BI gene e
	14	13.6	68.0	23	20	AAX24649	Human SR-BI gene e
	15	13.6	68.0	27	14	AAQ34929	Mutagenic beta glo
c	16	13.6	68.0	28	21	AAA63976	PCR primer used to
c	17	13.6	68.0	28	21	AAZ49806	Arabidopsis hydrop
	18	13.6	68.0	45	17	AAT38929	Primer 6 for const
	19	13.6	68.0	45	17	AAT39278	PCR primer LWN2043
	20	13.6	68.0	45	18	AAT75023	PCR primer 2 used
	21	13.6	68.0	45	19	AAV29510	SAVINASE gene and
c	22	13.6	68.0	48	19	AAV18213	Codon optimised si
c	23	13.6	68.0	60	15	AAQ67693	Codon-optimised D.
	24	13.6	68.0	65	24	ABN53455	Mouse spliced tran
c	25	13.6	68.0	74	15	AAQ67709	Fragment B cuticle
c	26	13.6	68.0	75	15	AAQ67332	AaIT optimised seq
c	27	13.6	68.0	75	19	AAV18225	Oligomer for a cod
c	28	13.6	68.0	75	24	ABK49981	Cuticle gene used
c	29	13.6	68.0	78	10	AAN90390	DNA encoding polyp
	30	13.2	66.0	33	20	AAX99344	Probe Oa-839R for
	31	13.2	66.0	43	19	AAV33646	Mouse VL30 promote
c	32	13.2	66.0	44	19	AAV57479	Cytochrome P450ox
	33	13.2	66.0	50	22	AAL34312	Human SNP oligonuc
	34	13.2	66.0	50	22	AAL34331	Human SNP oligonuc
	35	13.2	66.0	51	21	AAA77071	Human clone cg4392
c	36	13.2	66.0	51	22	AAL27598	Human SNP oligonuc
	37	13.2	66.0	51	22	AAI75224	Human silent SNP c
c	38	13.2	66.0	78	21	AAZ51490	Maize ZmKCS1 parti
	39	13	65.0	27	20	AAX56280	Human TAB1 PCR pri
	40	13	65.0	30	20	AAX99699	TAB1 cDNA amplifyi
	41	13	65.0	51	22	AAL29808	Human SNP oligonuc
	42	12.8	64.0	20	22	AAF31755	Human RANK antisen
c	43	12.8	64.0	21	21	AAZ76841	Human biallelic ma
c	44	12.8	64.0	24	24	ABI87616	Capture oligonucle
	45	12.8	64.0	24	24	ABI87617	Capture oligonucle

ALIGNMENTS

RESULT 1
AAZ31432
ID AAZ31432 standard; DNA; 20 BP.
XX
AC AAZ31432;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3602.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 16; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 5 A; 6 C; 9 G; 0 U; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGCACGGCGCAGAGGAGC 20
||||||||||||||||||||
Db 1 CGAGCACGGCGCAGAGGAGC 20

RESULT 2
ABK76729
ID ABK76729 standard; DNA; 99 BP.
XX
AC ABK76729;
XX
DT 13-AUG-2002 (first entry)
XX
DE Bacillus licheniformis genomic sequence tag (GST) #4020.
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus licheniformis.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US31437.
XX
PR 06-OCT-2000; 2000US-0680598.
PR 27-MAR-2001; 2001US-279526P.
XX
PA (NOVO) NOVOZYMES BIOTECH INC.
PA (NOVO) NOVOZYMES AS.
XX
PI Berka R, Clausen IG;
XX
DR WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second
PT Bacillus cells, by using substrate containing Bacillus genomic
PT sequenced tag array -
XX
PS Claim 4; SEQ ID NO 4020; 200pp; English.
XX
CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive
CC follow-up characterisation is unnecessary, when one spot on an array
CC equals one gene or one open reading frame, since sequence information is
CC available. This sequence represents a genomic sequence tag (GST) used in
CC the method of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 99 BP; 32 A; 26 C; 20 G; 20 T; 1 other;

Query Match 77.0%; Score 15.4; DB 24; Length 99;
Best Local Similarity 94.1%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GCACGGCGCAGAGGAGC 20
||||| ||||||||||
Db 31 GCACGCCGCAGAGGAGC 47

RESULT 3
ABK27825
ID ABK27825 standard; DNA; 27 BP.
XX
AC ABK27825;
XX
DT 09-APR-2002 (first entry)
XX
DE Corn male reproductive tissue promoter, PCR primer #15.
XX
KW Corn; male reproductive tissue; promoter; plant; gene stacking;
KW fertility; insect; pathogen; herbicide tolerance; primer; ss.
XX
OS Zea mays.
XX
PN WO200200905-A2.
XX
PD 03-JAN-2002.
XX
PF 26-JUN-2001; 2001WO-US20658.
XX
PR 28-JUN-2000; 2000US-214357P.
XX
PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
PI Conner TW, Dubois P, Malven M, Masucci JD;
XX
DR WPI; 2002-147890/19.
XX
PT New plant regulatory sequences or promoters and nucleic acids encoding
PT them, useful for regulating gene expression, especially in male
PT reproductive tissues -
XX
PS Example 3; Page 96; 131pp; English.
XX
CC The invention relates to an isolated nucleic acid (I) promoter capable of
CC regulating transcription of an operably linked DNA sequence. The promoter
CC sequences may be used for selectively modulating expression of any
CC operatively linked gene and provide additional regulatory element
CC diversity in a plant expression vector in gene stacking approaches, and
CC for regulating gene expression in male reproductive tissues or gene
CC transcription of any target gene (e.g. for controlling fertility, insect,
CC pathogen or herbicide tolerance). The nucleic acids are useful as
CC hybridisation probes or primers in hybridisation assays of other plant

```
CC   tissues to identify closely related or homologous genes and associated
CC   regulatory sequences. ABK27808-ABK27918 represent corn male
CC   reproductive tissue promoter sequences and PCR primers of the invention.
XX
SQ   Sequence 27 BP; 8 A; 6 C; 11 G; 2 T; 0 other;

  Query Match             76.0%;  Score 15.2;  DB 24;  Length 27;
  Best Local Similarity   85.0%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          |||| |||||| |||||| |
Db      7 CGAGGACGGCGAAGAGGATC 26


RESULT 4
AAN80520
ID   AAN80520 standard; DNA; 48 BP.
XX
AC   AAN80520;
XX
DT   03-OCT-2002  (updated)
DT   11-DEC-1990  (first entry)
XX
DE   Synthetic oligonucleotide used in prodn. of hybrid gene.
XX
KW   Synthetic oligonucleotide; hybrid gene; RESA 3'repeat;
KW   Plasmodium falciparum; malaria; vaccine.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..24
FT                   /*tag= a
FT                   /note="claimed peptide"
XX
PN   WO8802757-A.
XX
PD   21-APR-1988.
XX
PF   16-OCT-1987;   87WO-A000352.
XX
PR   17-OCT-1986;   86AU-0008551.
XX
PA   (SARA-) SARAMANE PTY. LTD.
XX
PI   Langford CJ,  Edwards SJ;
XX
DR   WPI; 1988-119462/17.
DR   P-PSDB; AAP80519.
XX
PT   New hybrid protein or polypeptide - contg. epitope(s) or mimotope(s) of
XX
PS   Disclosure; ; pp; English.
XX
CC   several antigens esp. useful as antimalarial antigens.
CC   The oligonucleotide encodes an epitope of a P.falciparum antigen.
CC   The peptide can be used in an antimalarial vaccine.
CC   (Updated on 03-OCT-2002 to add missing OS field.)
XX
SQ   Sequence 48 BP; 14 A; 12 C; 20 G; 2 T; 0 other;

  Query Match             74.0%;  Score 14.8;  DB 9;  Length 48;
  Best Local Similarity   88.9%;  Pred. No. 2.6e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 GAGCACGGCGCAGAGGAG 19
          ||||||| ||| ||||||
Db     13 GAGCACGACGCCGAGGAG 30


RESULT 5
AAV40985
ID   AAV40985 standard; DNA; 21 BP.
XX
AC   AAV40985;
XX
DT   25-SEP-1998  (first entry)
XX
DE   Primer TELAML1:1216L21 for abnormality detection.
XX
KW   PCR primer; chromosomal abnormality; abnormality detection; leukaemia;
KW   lymphoma; carcinoma; adenocarcinoma; sarcoma; glioma; neuroblastoma;
KW   medullablastoma; malignant melanoma; malignant neoplastic condition; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9824928-A2.
XX
PD   11-JUN-1998.
XX
PF   08-DEC-1997;   97WO-DK00556.
XX
PR   06-DEC-1996;   96DK-0001401.
XX
PA   (PALL/) PALLISGAARD N.
XX
PI   Hokland P,  Pallisgaard N;
XX
DR   WPI; 1998-333344/29.
XX
PT   Detection of chromosomal abnormalities - by subjecting patient
PT   sample nucleic acids to a multiplex molecular amplification
PT   procedure using primers specific for characteristic nucleic acid
PT   sequence
XX
PS   Claim 73; Page 85; 126pp; English.
XX
CC   This sequence represents a primer used in the method of the invention for
CC   the detection of the presence or absence of chromosomal abnormalities,
CC   each abnormality being associated with a condition in a subject and each
CC   being defined by at least one characteristic nucleic acid sequence. The
CC   method comprises: (a) obtaining a sample of nucleic acids derived from a
CC   subject which may harbour one of the chromosomal abnormalities;
CC   (b) subjecting the sample to a multiplex molecular amplification (MMA)
CC   procedure, where a number of the characteristic sequences, if present in
CC   a sufficient amount, will be amplified; (c) retrieving the product(s)
CC   from step (b), and detecting the presence and/or absence of an amplicon
CC   characteristic of the abnormal sequences to detect the presence or
CC   absence of corresponding chromosomal abnormalities; where the MMA
CC   procedure comprises the use of at least 7 mutually distinct primers (MDP)
CC   in one single reaction mixture, each of the primers defining an end of at
CC   least one characteristic nucleic acid sequence, and where at least one of
CC   the primers defines the first end of at least two characteristic nucleic
CC   acid sequences, the characteristic nucleic acid sequences each being
CC   determined in their opposite ends by MDP selected from the remainder of
CC   the MDP. The methods can be used for detecting chromosomal abnormalities
CC   associated with diseases including numerous leukaemia's, lymphoma's,
CC   carcinoma's, adenocarcinoma's, sarcoma's, glioma's, neuroblastoma's,
CC   medullablastoma, malignant melanoma, and malignant neoplastic conditions.
XX
SQ   Sequence 21 BP; 7 A; 3 C; 9 G; 2 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 19;  Length 21;
  Best Local Similarity   93.8%;  Pred. No. 4e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 AGCACGGCGCAGAGGA 18
          ||||||| ||||||||
Db      1 AGCACGGAGCAGAGGA 16


RESULT 6
ABK88264/c
ID   ABK88264 standard; DNA; 21 BP.
```

XX
AC ABK88264;
XX
DT 07-OCT-2002 (first entry)
XX
DE Exon 2 (alpha1 domain) of chicken class I gene PCR primer RV69BL.
XX
KW Chicken; exon 2; alpha1 domain; class I gene; Rfp-Y; B-F; B-L; B-G;
KW major histocompatibility complex; MHC; domesticated fowl;
KW Marek's disease; genetic disease resistance; PCR; primer; ss.
XX
OS Gallus sp.
XX
PN US2002083482-A1.
XX
PD 27-JUN-2002.
XX
PF 22-JUN-2001; 2001US-0886607.
XX
PR 19-JUL-2000; 2000US-0619758.
XX
PA (MILL/) MILLER M M.
PA (BRIL/) BRILES W E.
PA (AFAN/) AFANASSIEFF M.
XX
PI Miller MM, Briles WE, Afanassieff M;
XX
DR WPI; 2002-557335/59.
XX
PT Determining the major histocompatibility complex genotypes of a
PT domesticated fowl useful for selecting fowl, particularly chickens,
PT genetically resistant to a variety of diseases, particularly Marek's
PT disease -
XX
PS Claim 14; Page 7; 22pp; English.
XX
CC The present invention relates to a new method for determining the major
CC histocompatibility complex (MHC) genotypes of a domesticated fowl. The
CC method of the invention involves amplifying and denaturing nucleic acid
CC from the fowl which contains a region with a sequence corresponding to an
CC Rfp-Y, B-F, B-L or B-G region of the genome and separating the denatured
CC products electrophoretically to produce a pattern characteristic to the
CC genotype of the fowl. The method is useful for determining the MHC
CC genotypes of domestic fowl. The method may be used to select for domestic
CC fowl, particularly chickens, genetically resistant to a variety of
CC diseases, particularly Marek's disease. The methods of the invention are
CC useful for producing offspring that carry genetic resistance to disease.
CC The present nucleic acid sequence represents an exon 2 (alpha1 domain) of
CC chicken class I gene PCR primer that was used in the methods of the
CC invention for MHC genotyping.
XX
SQ Sequence 21 BP; 4 A; 7 C; 6 G; 4 T; 0 other;

Query Match 72.0%; Score 14.4; DB 24; Length 21;
Best Local Similarity 93.8%; Pred. No. 4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CACGGCGCAGAGGAGC 20
|||||||||||||||
Db 16 CACGGTGCAGAGGAGC 1

RESULT 7
AAD29196/c
ID AAD29196 standard; DNA; 21 BP.
XX
AC AAD29196;
XX
DT 07-MAY-2002 (first entry)
XX
DE Fowl MHc B class II-beta genotype analysing PCR primer, RV69BL.
XX
KW Fowl; haplotyping; breeding; domesticated fowl; disease resistance;
KW major histocompatibility; MHC; Marek's disease; PCR primer; ss.
XX
OS Gallus sp.
XX
PN WO200206535-A2.
XX
PD 24-JAN-2002.
XX
PF 19-JUL-2001; 2001WO-US22519.
XX
PR 19-JUL-2000; 2000US-0619758.
XX
PA (CITY) CITY OF HOPE.
XX
PI Miller MM, Afanassieff M, Briles WE;
XX
DR WPI; 2002-171822/22.
XX
PT Determining major histocompatibility complex genotypes of domesticated
PT fowl for breeding disease resistant fowl by comparing B and Rfp-Y
PT systems of genome from the fowl to corresponding known genotypes -
XX
PS Claim 14; Page 16; 37pp; English.
XX
CC The invention relates to a method for haplotyping and breeding
CC domesticated fowl for increased disease resistance. Disease
CC resistance in domesticated fowl has been associated with the B and
CC Rfp-Y systems of major histocompatibility (MHC) genes. The method
CC is useful for determining the Mhc genotypes of a domesticated fowl.
CC The method is useful in selecting domesticated fowl which are
CC resistant to a preselected disease preferably Marek's disease by
CC determining a genotype of Rfp-Y, B-F, B-L or B-G genes of the
CC domesticated fowl, correlating the determined genotypes with
CC resistance to the preselected disease and selecting a domesticated
CC fowl having a determined genotype which correlates with resistance
CC to the preselected disease. The method is also useful for breeding
CC domesticated fowl to produce an offspring resistant to a preselected
CC disease by selecting a first domesticated fowl of a gender which is
CC resistant to a preselected disease, and mating the first domesticated
CC fowl with a second domesticated fowl of the opposite gender to produce
CC offspring. The present sequence is a PCR primer used for genotyping
CC fowl Mhc region.
XX
SQ Sequence 21 BP; 4 A; 7 C; 6 G; 4 T; 0 other;

Query Match 72.0%; Score 14.4; DB 24; Length 21;
Best Local Similarity 93.8%; Pred. No. 4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CACGGCGCAGAGGAGC 20
|||||||||||||||
Db 16 CACGGTGCAGAGGAGC 1

RESULT 8
AAI73383
ID AAI73383 standard; DNA; 50 BP.
XX
AC AAI73383;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ:324.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.

```
XX
PF   30-NOV-2000; 2000WO-US32758.
XX
PR   30-NOV-1999;   99US-0168138.
PR   29-NOV-2000; 2000US-0726173.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-356160/37.
XX
PT   Polymorphic nucleic acid sequences, useful in genetic testing and
PT   therapy -
XX
PS   Claim 1; Page 153; 2653pp; English.
XX
CC   AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC   sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC   AAM53114 to AAM53329 represent peptides related to human polymorphic
CC   polynucleotide sequences. The sequences can be used in gene and protein
CC   therapy, and in vaccine production. (I) and the polypeptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of polymorphic polypeptides.
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 50 BP; 5 A; 14 C; 28 G; 3 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 50;
  Best Local Similarity   84.2%;  Pred. No. 4.7e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAG 19
           |||| ||||||| || |||
Db      30 CGAGGACGGCGCGGACGAG 48


RESULT 9
AAA77070
ID   AAA77070 standard; cDNA; 51 BP.
XX
AC   AAA77070;
XX
DT   16-NOV-2000  (first entry)
XX
DE   Human clone cg43921619 polymorphic site, SEQ ID NO:753.
XX
KW   Human; single nucleotide polymorphism; SNP; chromosome 17;
KW   detection; identification; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace (26,T)
FT                   /*tag= a
XX
PN   WO200029623-A2.
XX
PD   25-MAY-2000.
XX
PF   17-NOV-1999;   99WO-US27293.
XX
PR   17-NOV-1998;   98US-0109024.
PR   16-NOV-1999;   99US-0109024.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach MD;
XX
DR   WPI; 2000-387826/33.
XX
PT   Human nucleic acids containing single nucleotide polymorphisms, useful
PT   for treating a subject suffering, or at risk from a pathology due to
PT   the presence of a sequence polymorphism -
XX
PS   Claim 1; Page 385; 543pp; English.
XX
CC   Sequences AAA76318-A77509 represent 1192 human nucleic acid sequences
CC   which contain single nucleotide polymorphisms (SNPs). Sequences 1 to
CC   1112 (AAA76318-A77429) are consecutive pairs of nucleotides which
CC   contain silent SNPs. Sequences 1113 to 1192 (AAA77430-A77509) are
CC   consecutive pairs of nucleotides containing SNPs which result in changes
CC   in the corresponding amino acid sequences (AAB11749-B11828). The SNPs in
CC   sequences 1113 to 1128 (AAA77430-A77445) lead to conservative amino acid
CC   changes, while those in sequences 1129 to 1186 (AAA77446-A77503) result
CC   in non- conservative changes. The SNPs in sequences 1187 to 1192
CC   (AAA77504-A77509) generate frameshift mutations. The invention also
CC   relates to a method of detecting a polymorphic site in a nucleic acid and
CC   a method of determining the relatedness of two nucleic acids. It also
CC   encompasses peptides containing polymorphic sites, antibodies raised
CC   against such peptides, and a method of detecting polymorphic
CC   proteins/peptides using the antibodies. The nucleic acids are useful for
CC   gene therapy of an individual having, suspected of having, or at risk of
CC   developing a pathological condition due to the presence of a sequence
CC   polymorphism. Such treatment would comprise administration of the
CC   wild-type nucleic acid sequence. Antibodies raised against polymorphic
CC   peptides can also be used in the treatment of such individuals.
XX
SQ   Sequence 51 BP; 15 A; 4 C; 24 G; 8 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 21;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 4.7e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGCACGGCGCAGAGGAGC 20
           ||||| || | ||||||||
Db      26 GAGCAGGGAGAAGAGGAGC 44


RESULT 10
AAI73382
ID   AAI73382 standard; DNA; 51 BP.
XX
AC   AAI73382;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:323.
XX
KW   Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW   protein therapy; vaccine; probe; diagnostic assay; detection;
KW   quantitation; restorative therapy; polymorphic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200140521-A2.
XX
PD   07-JUN-2001.
XX
PF   30-NOV-2000; 2000WO-US32758.
XX
PR   30-NOV-1999;   99US-0168138.
PR   29-NOV-2000; 2000US-0726173.
XX
PA   (CURA-) CURAGEN CORP.
```

```
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-356160/37.
XX
PT   Polymorphic nucleic acid sequences, useful in genetic testing and
PT   therapy -
XX
PS   Claim 1; Page 153; 2653pp; English.
XX
CC   AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC   sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC   AAM53114 to AAM53329 represent peptides related to human polymorphic
CC   polynucleotide sequences. The sequences can be used in gene and protein
CC   therapy, and in vaccine production. (I) and the polypeptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of polymorphic polypeptides.
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 51 BP; 5 A; 14 C; 29 G; 3 T; 0 other;

  Query Match           71.0%;  Score 14.2;  DB 22;  Length 51;
  Best Local Similarity  84.2%;  Pred. No. 4.7e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAG 19
           |||| ||||||| || |||
Db      31 CGAGGACGGCGCGGACGAG 49


RESULT 11
AAZ57918/c
ID   AAZ57918 standard; DNA; 37 BP.
XX
AC   AAZ57918;
XX
DT   11-APR-2000  (first entry)
XX
DE   Escherichia coli trpC gene 3' flanking region PCR primer N19.
XX
KW   TrpC gene; N-(5'-phosphoribosyl)-anthranilate isomerase;
KW   expression system; selectable marker; auxotroph complementation;
KW   PCR primer; ss.
XX
OS   Escherichia coli.
XX
PN   EP972838-A1.
XX
PD   19-JAN-2000.
XX
PF   08-JUL-1999;   99EP-0113182.
XX
PR   15-JUL-1998;   98EP-0113156.
PR   09-OCT-1998;   98EP-0119078.
XX
PA   (HOFF ) ROCHE DIAGNOSTICS GMBH.
XX
PI   Kopetzki E,  Schantz C;
XX
DR   WPI; 2000-108129/10.
XX
PT   Novel prokaryotic expression system, used to produce recombinant
PT   proteins -
XX
PS   Example 6; Page 13; 31pp; English.
XX
CC   The present sequence is that of primer N19, which was used as
CC   antisense primer in the PCR amplification of the 3' flanking region
CC   of the Escherichia coli trpC gene (trpB) using chromosomal DNA as
CC   template. The primer introduces a 3' SphI cleavage site into the
CC   amplified product. The PCR product was used in the construction of
CC   deletion plasmid pMAK705-trpD/B. A trpC auxotrophy marker was
CC   introduced into a laboratory E. coli strain by directed deletion
CC   via homologous recombination in order to create a non-revertible
CC   trpC host characterised by an inactive N-(5'-phosphoribosyl)-
CC   anthranilate isomerase. This host strain can be used in a novel
CC   host/vector system of the invention. The invention provides a
CC   stable prokaryotic (preferably Escherichia coli) host/vector system
CC   (expression system) which avoids plasmid selection by antibiotics.
CC   The selection principle is based on complementation of a stable
CC   auxotrophy (e.g. trpC, pyrF) of the host cells by a yeast gene (e.g.
CC   TRP1, URA3). A minimal prokaryotic expression vector that cannot be
CC   homologously recombined with the genome of prokaryotic organisms
CC   includes an origin of replication, an auxotrophy marker gene, a
CC   promoter that is functionally active in prokaryotes; and a foreign
CC   gene to be expressed under the control of the promoter. The new
CC   expression system is used for the production of recombinant proteins.
XX
SQ   Sequence 37 BP; 8 A; 11 C; 11 G; 7 T; 0 other;

  Query Match           69.0%;  Score 13.8;  DB 21;  Length 37;
  Best Local Similarity  88.2%;  Pred. No. 7e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGG 17
           |||||||||||||  ||
Db      28 CGAGCACGGCGCACCGG 12


RESULT 12
ABN56998
ID   ABN56998 standard; DNA; 65 BP.
XX
AC   ABN56998;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse spliced transcript detection oligonucleotide SEQ ID NO:29746.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Mus musculus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 29746; 47pp; English.
```

```
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 17 A; 17 C; 24 G; 7 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 65;
  Best Local Similarity   88.2%;  Pred. No. 6.9e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GAGCACGGCGCAGAGGA 18
           ||||| ||||||| |||
Db      37 GAGCATGGCGCAGGGGA 53


RESULT 13
AAX24557
ID   AAX24557 standard; DNA; 23 BP.
XX
AC   AAX24557;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Human SR-BI gene exon 1 PCR primer.
XX
KW   SR-BI; human; polymorphism; cardiovascular disorder; ischaemia;
KW   restenosis; congestive heart failure; atherosclerosis; cholesterol;
KW   low density lipoprotein; LDL; high density lipoprotein; HDL;
KW   diagnosis; body mass index; obesity; cachexia; gallstone; PCR;
KW   primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9902735-A2.
XX
PD   21-JAN-1999.
XX
PF   10-JUL-1998;   98WO-US14354.
XX
PR   27-FEB-1998;   98US-0031626.
PR   10-JUL-1997;   97US-0890979.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
PA   (UYTU-) UNIV TUFTS.
XX
PI   Acton SL,  Ordovas JM;
XX
DR   WPI; 1999-120935/10.
XX
PT   Detecting genetic predisposition for body mass disorders - by
PT   identifying allelic variants of a polymorphic region of the SR-BI
PT   gene
XX
PS   Example 5; Page 72; 102pp; English.
XX
CC   A PCR primer pair (see also AAX24556) is designed for the
CC   amplification of exon 1 (see AAX24498) of the human SR-BI gene.
CC   A G/A polymorphism has been detected at nucleotide 146 of this
CC   exon.  PCR amplification followed by AluI digestion yields
CC   a 263 bp product in GG individuals, 263, 192 and 71 bp products
CC   in GA individuals, and 192 and 71 bp products in AA individuals.
CC   The invention is based on the discovery of the genomic structure of
CC   the human SR-BI gene (see AAX24498-509) and on the identification of
CC   polymorphic regions within the gene which are associated with
CC   abnormal body mass index (BMI) and abnormal lipoprotein levels and
CC   hence with disorders such as obesity, cachexia, cardiovascular
CC   disorders and gallstone formation.  The invention provides methods
CC   for determining whether a subject has, or is at risk of developing,
CC   a disease associated with a specific allele of a polymorphic region
CC   of an SR-BI gene.  Kits comprising the relevant probe or primer are
CC   claimed.
XX
SQ   Sequence 23 BP; 6 A; 10 C; 6 G; 1 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 20;  Length 23;
  Best Local Similarity   80.0%;  Pred. No. 8.6e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           | ||||| ||||| || |||
Db       4 CCAGCACAGCGCACAGTAGC 23


RESULT 14
AAX24649
ID   AAX24649 standard; DNA; 23 BP.
XX
AC   AAX24649;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Human SR-BI gene exon 1 PCR primer.
XX
KW   SR-BI; human; polymorphism; cardiovascular disorder; ischaemia;
KW   restenosis; congestive heart failure; atherosclerosis; cholesterol;
KW   low density lipoprotein; LDL; high density lipoprotein; HDL;
KW   diagnosis; body mass index; obesity; cachexia; gallstone; PCR;
KW   primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9902736-A2.
XX
PD   21-JAN-1999.
XX
PF   10-JUL-1998;   98WO-US14359.
XX
PR   27-FEB-1998;   98US-0032894.
PR   10-JUL-1997;   97US-0890980.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Acton SL;
XX
DR   WPI; 1999-120936/10.
XX
PT   New nucleic acids comprising intronic sequence of a human scavenger
PT   receptor-BI (SR-B1) gene  - useful for prognosis, diagnosis and
PT   treatment of SR-BI associated diseases or conditions
XX
PS   Claim 10; Page 71; 103pp; English.
XX
```

CC A PCR primer pair (see also AAX24648) is designed for the
CC amplification of exon 1 (see AAX24590) of the human SR-BI gene.
CC A G/A polymorphism has been detected at nucleotide 146 of this
CC exon. PCR amplification followed by AluI digestion yields
CC a 263 bp product in GG individuals, 263, 192 and 71 bp products
CC in GA individuals, and 192 and 71 bp products in AA individuals.
CC The invention is based on the discovery of the genomic structure of
CC the human SR-BI gene (see AAX24590-601) and on the identification of
CC polymorphic regions within the gene which are associated with
CC abnormal body mass index (BMI) and abnormal lipoprotein levels and
CC hence with disorders such as obesity, cachexia, cardiovascular
CC disorders and gallstone formation. The invention provides methods
CC for determining whether a subject has, or is at risk of developing,
CC a disease associated with a specific allele of a polymorphic region
CC of an SR-BI gene. Kits comprising the relevant probe or primer are
CC claimed.
XX
SQ Sequence 23 BP; 6 A; 10 C; 6 G; 1 T; 0 other;

Query Match 68.0%; Score 13.6; DB 20; Length 23;
Best Local Similarity 80.0%; Pred. No. 8.6e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGAGCACGGCGCAGAGGAGC 20
| ||||| ||||| || |||
Db 4 CCAGCACAGCGCACAGTAGC 23

RESULT 15
AAQ34929
ID AAQ34929 standard; DNA; 27 BP.
XX
AC AAQ34929;
XX
DT 13-MAY-1993 (first entry)
XX
DE Mutagenic beta globin primer #3.
XX
KW Globin; haemoglobin; Hb; affinity; oxygen; delivery; stroke;
KW myocardial infarction; diabetes; tumour; X-ray treatment; burns;
KW free radical; lesion; PCR; polymerase chain reaction; amplify; ss.
XX
OS Synthetic.
XX
PN US5173426-A.
XX
PD 22-DEC-1992.
XX
PF 06-OCT-1989; 89US-0417949.
XX
PR 06-OCT-1989; 89US-0417949.
XX
PA (UYYA) UNIV YALE.
XX
PI Baserga SJ, Fischer JJ;
XX
DR WPI; 1993-017570/02.
XX
PT New polynucleotide(s) coding for mutant globin chains - form
PT haemoglobin with reduced oxygen affinity, useful e.g. for
PT increasing oxygenation of tumour cells
XX
PS Disclosure; Column 14; 11pp; English.
XX
CC The sequences given in AAQ34927-30 are primers which were used in the
CC production of mutant globin molecules which when part of a haemoglobin
CC (Hb) molecule, produces a mutant Hb with a lower affinity for oxygen
CC than normal Hb. The oxygen affinity of the mutant Hb's being
CC characterised by P50 of 30 Torr to 3 atmos. and/or Hill coefficient
CC of 2.5-1. The mutant Hb are useful for increasing oxygen delivery to
CC tissue in patients suffering from myocardial infarction, stroke,
CC diabetes etc. They may also be used to make tumours more susceptible
CC to X-ray treatment by combining with free radicals and fixing the
CC lesion in the cell. These Hb may be used replace the Hb in an entire
CC bloodstream and for treating burns patients.
XX
SQ Sequence 27 BP; 7 A; 11 C; 7 G; 2 T; 0 other;

Query Match 68.0%; Score 13.6; DB 14; Length 27;
Best Local Similarity 80.0%; Pred. No. 8.6e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGAGCACGGCGCAGAGGAGC 20
| |||||| ||| ||||||
Db 6 CCAGCACGTCGCCCAGGAGC 25

Search completed: November 23, 2002, 06:28:30
Job time : 103.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-2
Perfect score: 20
Sequence: 1 cgagcacggcgcagaggagc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	13.8	69.0	37	4	US-09-344-888A-19	Sequence 19, Appl
	2	13.6	68.0	23	3	US-09-032-894-90	Sequence 90, Appl
	3	13.6	68.0	23	4	US-09-031-626-90	Sequence 90, Appl
	4	13.6	68.0	27	1	US-08-170-095B-36	Sequence 36, Appl
	5	13.6	68.0	27	6	5428007-6	Patent No. 5428007
	6	13.6	68.0	45	1	US-08-785-448-6	Sequence 6, Appli
	7	13.6	68.0	45	2	US-08-925-927-6	Sequence 6, Appli
	8	13.6	68.0	45	2	US-08-875-154-8	Sequence 8, Appli
	9	13.6	68.0	45	3	US-09-120-576-6	Sequence 6, Appli
c	10	13.6	68.0	60	1	US-08-009-265-17	Sequence 17, Appl
c	11	13.6	68.0	75	1	US-08-009-265-42	Sequence 42, Appl
c	12	13.6	68.0	75	2	US-08-322-679-10	Sequence 10, Appl
c	13	13.6	68.0	75	4	US-08-686-993A-22	Sequence 22, Appl
c	14	13.6	68.0	78	6	5457037-7	Patent No. 5457037
	15	13.2	66.0	43	4	US-09-380-190A-46	Sequence 46, Appl
c	16	13.2	66.0	44	4	US-09-380-420C-13	Sequence 13, Appl
	17	13.2	66.0	92	1	US-08-487-141B-109	Sequence 109, App
	18	13	65.0	27	4	US-09-529-279-7	Sequence 7, Appli
	19	13	65.0	37	1	US-08-253-785-12	Sequence 12, Appl
c	20	13	65.0	45	1	US-08-253-785-13	Sequence 13, Appl
	21	12.8	64.0	20	4	US-09-435-296-13	Sequence 13, Appl
	22	12.8	64.0	75	2	US-09-073-032-4	Sequence 4, Appli
c	23	12.6	63.0	30	3	US-08-929-856-63	Sequence 63, Appl
c	24	12.6	63.0	33	4	US-09-349-627-3	Sequence 3, Appli
	25	12.6	63.0	36	3	US-08-767-942A-30	Sequence 30, Appl
	26	12.6	63.0	52	1	US-08-685-793-13	Sequence 13, Appl
	27	12.6	63.0	52	1	US-08-685-793-20	Sequence 20, Appl
	28	12.6	63.0	53	1	US-08-434-743-3	Sequence 3, Appli
	29	12.6	63.0	53	1	US-08-685-793-5	Sequence 5, Appli
	30	12.6	63.0	56	1	US-08-685-793-9	Sequence 9, Appli
c	31	12.6	63.0	60	1	US-08-349-696-4	Sequence 4, Appli
c	32	12.6	63.0	60	1	US-08-233-009-4	Sequence 4, Appli
c	33	12.6	63.0	60	1	US-08-560-231-4	Sequence 4, Appli
c	34	12.6	63.0	60	4	US-09-080-704A-4	Sequence 4, Appli
	35	12.6	63.0	61	3	US-09-023-228B-110	Sequence 110, App
	36	12.6	63.0	61	4	US-09-163-025B-110	Sequence 110, App
	37	12.6	63.0	99	1	US-08-472-255A-50	Sequence 50, Appl
	38	12.6	63.0	99	1	US-08-479-724A-50	Sequence 50, Appl
	39	12.6	63.0	99	3	US-08-472-256B-50	Sequence 50, Appl
	40	12.6	63.0	99	4	US-08-952-793-50	Sequence 50, Appl
	41	12.6	63.0	99	5	PCT-US96-09455A-50	Sequence 50, Appl
c	42	12.4	62.0	20	1	US-08-468-447-12	Sequence 12, Appl
c	43	12.4	62.0	20	1	US-08-469-851A-12	Sequence 12, Appl
c	44	12.4	62.0	20	1	US-08-467-597A-12	Sequence 12, Appl
c	45	12.4	62.0	20	1	US-08-468-569A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-344-888A-19/c
; Sequence 19, Application US/09344888A
; Patent No. 6291245
; GENERAL INFORMATION:
; APPLICANT: Kopetzki, Erhard
; APPLICANT: Schantz, Christian
; TITLE OF INVENTION: New Host-Vector System
; FILE REFERENCE: CD20315
; CURRENT APPLICATION NUMBER: US/09/344,888A
; CURRENT FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: EP98113156.8
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: EP98119078.8
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-344-888A-19

Query Match 69.0%; Score 13.8; DB 4; Length 37;
Best Local Similarity 88.2%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGAGCACGGCGCAGAGG 17
||||||||||||| ||
Db 28 CGAGCACGGCGCACCGG 12

RESULT 2
US-09-032-894-90
; Sequence 90, Application US/09032894
; Patent No. 6130041
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: SR-BI NUCLEIC ACIDS AND USES THEREFOR
; FILE REFERENCE: MIA-005.03
; CURRENT APPLICATION NUMBER: US/09/032,894
; CURRENT FILING DATE: 1998-02-27
; EARLIER APPLICATION NUMBER: 08/890,980
; EARLIER FILING DATE: 1997-07-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 23

```
;   TYPE: DNA
;   ORGANISM: Human
US-09-032-894-90

  Query Match             68.0%;  Score 13.6;  DB 3;  Length 23;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          | ||||| ||||| || |||
Db      4 CCAGCACAGCGCACAGTAGC 23


RESULT 3
US-09-031-626-90
; Sequence 90, Application US/09031626
; Patent No. 6228581
; GENERAL INFORMATION:
;  APPLICANT: Acton, Susan L.
;  APPLICANT: Ordovas, Jose M.
;  TITLE OF INVENTION: DIAGNOSTIC ASSAYS AND KITS FOR BODY MASS AND
;  TITLE OF INVENTION: CARDIOVASCULAR DISORDERS
;  FILE REFERENCE: MIA-005.04
;  CURRENT APPLICATION NUMBER: US/09/031,626
;  CURRENT FILING DATE: 1998-02-27
;  EARLIER APPLICATION NUMBER: 08/890,979
;  EARLIER FILING DATE: 1997-07-10
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
;   LENGTH: 23
;   TYPE: DNA
;   ORGANISM: Human
US-09-031-626-90

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 23;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          | ||||| ||||| || |||
Db      4 CCAGCACAGCGCACAGTAGC 23


RESULT 4
US-08-170-095B-36
; Sequence 36, Application US/08170095B
; Patent No. 5563254
;  GENERAL INFORMATION:
;    APPLICANT:  Hoffman, Stephen J.
;    APPLICANT:  Nagai, Kiyoshi
;    TITLE OF INVENTION:  Blood Substitutes
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Somatogen, Inc.
;      STREET:  2545 Central Avenue
;      CITY:  Boulder
;      STATE:  Colorado
;      ZIP:  80301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4 Mb storage
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  System 7.0.1
;      SOFTWARE:  Microsoft Word 5.0a
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/170,095B
;      FILING DATE:  December 20, 1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5563254ak, Henry P.
;      REGISTRATION NUMBER:  33200
;      REFERENCE/DOCKET NUMBER:  Hoffman 2A/CONT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303-541-3322
;      TELEFAX:  303-444-3013
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown to applicant
;    MOLECULE TYPE:  Other nucleic acid
;      DESCRIPTION:  beta globin sequence fragment
;    HYPOTHETICAL:  no
US-08-170-095B-36

  Query Match             68.0%;  Score 13.6;  DB 1;  Length 27;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          | |||||| |||  ||||||
Db      6 CCAGCACGTCGCCCAGGAGC 25


RESULT 5
5428007-6
;Patent No. 5428007
;    APPLICANT: FISCHER, JAMES J.;BASERGA, SUSAN J.
;    TITLE OF INVENTION: GENETICALLY ENGINEERED LOW OXYGEN
;AFFINITY MUTANTS OF HUMAN HEMOGLOBIN
;    NUMBER OF SEQUENCES: 7
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/235,118
;      FILING DATE: 28-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 959,286
;      FILING DATE: 09-OCT-1992
;      APPLICATION NUMBER: 417,949
;      FILING DATE: 06-OCT-1989
;SEQ ID NO:6:
;      LENGTH: 27
5428007-6

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 27;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          | |||||| |||  ||||||
Db      6 CCAGCACGTCGCCCAGGAGC 25


RESULT 6
US-08-785-448-6
; Sequence 6, Application US/08785448
; Patent No. 5763187
;  GENERAL INFORMATION:
;    APPLICANT:  Sloma, Alan P.
;    APPLICANT:  Widner, William R.
;    TITLE OF INVENTION:  Bacterial Donor Cell Useful
;    TITLE OF INVENTION:  In Conjugation
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5763187o No. 5763187disk of No. 5763187th America
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/785,448
;      FILING DATE:  17-JAN-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Agris, Cheryl H
;      REGISTRATION NUMBER:  34,086
;      REFERENCE/DOCKET NUMBER:  4666.200-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-785-448-6

  Query Match             68.0%;  Score 13.6;  DB 1;  Length 45;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           || || ||||| |||||| |
Db      16 CGGGCCCGGCGTAGAGGATC 35


RESULT 7
US-08-925-927-6
; Sequence 6, Application US/08925927
; Patent No. 5843720
;  GENERAL INFORMATION:
;    APPLICANT:  Tangney, Martin
;    APPLICANT:  Hansen, Christian
;    APPLICANT:  Pedersen, Poul Erik
;    APPLICANT:  Jorgensen, Per Lina
;    APPLICANT:  Jorgensen, Steen Troels
;    TITLE OF INVENTION:  Introduction Of DNA Into Bacillus
;    TITLE OF INVENTION:  By Conjugation
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5843720o No. 5843720disk of No. 5843720th America
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/925,927
;      FILING DATE:  08-SEP-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4125.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-925-927-6

  Query Match             68.0%;  Score 13.6;  DB 2;  Length 45;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           || || ||||| |||||| |
Db      16 CGGGCCCGGCGTAGAGGATC 35


RESULT 8
US-08-875-154-8
; Sequence 8, Application US/08875154
; Patent No. 5882888
;  GENERAL INFORMATION:
;    APPLICANT:  Jorgensen, Streen Troels
;    TITLE OF INVENTION:  DNA Integration By Transporation
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5882888o No. 5882888disk of No. 5882888th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/875,154
;      FILING DATE:  17-JUL-1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4381.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "LWN2043"
US-08-875-154-8

  Query Match             68.0%;  Score 13.6;  DB 2;  Length 45;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           || || ||||| |||||| |
Db      16 CGGGCCCGGCGTAGAGGATC 35


RESULT 9
US-09-120-576-6
; Sequence 6, Application US/09120576
; Patent No. 6066473
;  GENERAL INFORMATION:
;    APPLICANT:  Tangney, Martin
;    APPLICANT:  Hansen, Christian
;    APPLICANT:  Pedersen, Poul Erik
;    APPLICANT:  Jorgensen, Per Lina
;    APPLICANT:  Jorgensen, Steen Troels
;    TITLE OF INVENTION:  Introduction Of DNA Into Bacillus
;    TITLE OF INVENTION:  By Conjugation
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  No. 6066473o No. 6066473disk of No. 6066473th America
;       STREET:  405 Lexington Avenue
;       CITY:  New York
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10174
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/120,576
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/925,927
;       FILING DATE:  08-SEP-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lambiris, Elias J
;       REGISTRATION NUMBER:  33,728
;       REFERENCE/DOCKET NUMBER:  4125.204-US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-867-0123
;       TELEFAX:  212-878-9655
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  45 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-09-120-576-6

  Query Match             68.0%;  Score 13.6;  DB 3;  Length 45;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           || || ||||| |||||| |
Db      16 CGGGCCCGGCGTAGAGGATC 35


RESULT 10
US-08-009-265-17/c
; Sequence 17, Application US/08009265
; Patent No. 5547871
;  GENERAL INFORMATION:
;    APPLICANT:  Black Dr., Bruce C.
;    APPLICANT:  Summers Dr., Max D.
;    TITLE OF INVENTION:  Heterologous Signal Sequences For
;    TITLE OF INVENTION:  Secretion Of Insect Controlling Proteins
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  American Cyanamid Company
;      STREET:  1937 West Main Street, P.O. Box 60
;      CITY:  Stamford
;      STATE:  CT
;      COUNTRY:  USA
;      ZIP:  06904-0060
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/009,265
;      FILING DATE:  19930125
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gordon Mr., Alan M.
;      REGISTRATION NUMBER:  30637
;      REFERENCE/DOCKET NUMBER:  31868-00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  203-321-2719
;      TELEFAX:  203-321-2971
;      TELEX:  710-474-4059
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila melanogaster
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  13..60
US-08-009-265-17

  Query Match             68.0%;  Score 13.6;  DB 1;  Length 60;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           |||| ||||||||||  | |
Db      46 CGAGGACGGCGCAGATCATC 27


RESULT 11
US-08-009-265-42/c
; Sequence 42, Application US/08009265
; Patent No. 5547871
;  GENERAL INFORMATION:
;    APPLICANT:  Black Dr., Bruce C.
;    APPLICANT:  Summers Dr., Max D.
;    TITLE OF INVENTION:  Heterologous Signal Sequences For
;    TITLE OF INVENTION:  Secretion Of Insect Controlling Proteins
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  American Cyanamid Company
;      STREET:  1937 West Main Street, P.O. Box 60
;      CITY:  Stamford
;      STATE:  CT
;      COUNTRY:  USA
;      ZIP:  06904-0060
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/009,265
;      FILING DATE:  19930125
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gordon Mr., Alan M.
;      REGISTRATION NUMBER:  30637
;      REFERENCE/DOCKET NUMBER:  31868-00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  203-321-2719
;      TELEFAX:  203-321-2971
;      TELEX:  710-474-4059
;  INFORMATION FOR SEQ ID NO:  42:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  75 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
```

```
;      ORGANISM:  Drosophila melanogaster
US-08-009-265-42

  Query Match            68.0%;  Score 13.6;  DB 1;  Length 75;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          |||| ||||||||||  | |
Db     46 CGAGGACGGCGCAGATCATC 27


RESULT 12
US-08-322-679-10/c
; Sequence 10, Application US/08322679
; Patent No. 5965123
;  GENERAL INFORMATION:
;    APPLICANT:  Ahmed, Fakhruddin
;    TITLE OF INVENTION:  Coated Pesticidal Agents, Processes For
;    TITLE OF INVENTION:  Their Preparation And Compositions Containing Them
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  American Cyanamid Company
;      STREET:  One Cyanamid Plaza
;      CITY:  Wayne
;      STATE:  New Jersey
;      COUNTRY:  United States
;      ZIP:  07470
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/322,679
;      FILING DATE:  13-OCT-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hogan, John W.
;      REGISTRATION NUMBER:  32,703
;      REFERENCE/DOCKET NUMBER:  32,458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-831-3195
;      TELEFAX:  201-831-3305
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  75 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-322-679-10

  Query Match            68.0%;  Score 13.6;  DB 2;  Length 75;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          |||| ||||||||||  | |
Db     46 CGAGGACGGCGCAGATCATC 27


RESULT 13
US-08-686-993A-22/c
; Sequence 22, Application US/08686993A
; Patent No. 6355240
; GENERAL INFORMATION:
;  APPLICANT:  Dierks, Peter
;  TITLE OF INVENTION:  ENHANCED INSECTICIDAL INSECT VIRUS THROUGH THE
;  TITLE OF INVENTION:  EXPRESSION OF HETEROLOGOUS PROTEINS WITH EARLY
;  TITLE OF INVENTION:  PROMOTERS
;  FILE REFERENCE:  0646/0D616
;  CURRENT APPLICATION NUMBER: US/08/686,993A
;  CURRENT FILING DATE:  1996-07-25
;  NUMBER OF SEQ ID NOS:  35
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 22
;   LENGTH: 75
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: cDNA
US-08-686-993A-22

  Query Match            68.0%;  Score 13.6;  DB 4;  Length 75;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          |||| ||||||||||  | |
Db     46 CGAGGACGGCGCAGATCATC 27


RESULT 14
5457037-7/c
;Patent No. 5457037
;    APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;
;GALLI, GIULIANO;LUCCHESE, GIUSEPPE;GRANDI, GUIDO;DI GENNARO, CARLO
;    TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE
; ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME
;    NUMBER OF SEQUENCES: 7
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/1,797
;      FILING DATE: 08-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 749,621
;      FILING DATE: 19-AUG-1991
;      APPLICATION NUMBER: 224,114
;      FILING DATE: 25-JUL-1988
;SEQ ID NO:7:
;      LENGTH: 78
5457037-7

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 78;
  Best Local Similarity   80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          ||||||| ||||||   |||
Db     49 CGAGCACCGCGCAGCCAAGC 30


RESULT 15
US-09-380-190A-46
; Sequence 46, Application US/09380190A
; Patent No. 6410220
;   GENERAL INFORMATION:
;        APPLICANT: NATURE TECHNOLOGY CORPORATION, ET AL.
;        TITLE OF INVENTION: SELF-ASSEMBLING GENES, VECTORS AND USES
;                            THEREOF
;        NUMBER OF SEQUENCES: 74
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
;             STREET: 119 NORTH FOURTH STREET, SUITE 203
;             CITY: MINNEAPOLIS
;             STATE: MINNESOTA
;             COUNTRY: USA
;             ZIP: 55401
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
```

```
;             APPLICATION NUMBER: US/09/380,190A
;             FILING DATE: 26-Aug-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/US98/03918
;             FILING DATE: 28-FEB-98
;        ATTORNEY/AGENT INFORMATION:
;             NAME: MUETING, ANN M.
;             REGISTRATION NUMBER: 33,977
;             REFERENCE/DOCKET NUMBER: 228.00010201
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 612-305-1217
;             TELEFAX: 612-305-1228
;   INFORMATION FOR SEQ ID NO: 46:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 43 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-380-190A-46

  Query Match            66.0%;  Score 13.2;  DB 4;  Length 43;
  Best Local Similarity  83.3%;  Pred. No. 2.1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GAGCACGGCGCAGAGGAG 19
           |||||| | | |||||||
Db      13 GAGCACAGGGTAGAGGAG 30


Search completed: November 23, 2002, 06:35:58
Job time : 24.55 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-2
Perfect score: 20
Sequence: 1 cgagcacggcgcagaggagc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	15.4	77.0	99	10	US-09-974-300-4020	Sequence 4020, Ap
	2	15.2	76.0	27	10	US-09-894-633A-18	Sequence 18, Appl
c	3	14.4	72.0	21	10	US-09-886-607-14	Sequence 14, Appl
	4	13	65.0	27	9	US-10-158-895-7	Sequence 7, Appli
	5	12.8	64.0	75	10	US-09-972-809-4	Sequence 4, Appli
	6	12.8	64.0	88	10	US-09-864-761-19378	Sequence 19378, A
	7	12.6	63.0	25	10	US-09-866-108-4429	Sequence 4429, Ap
	8	12.6	63.0	25	10	US-09-866-108-4430	Sequence 4430, Ap
	9	12.6	63.0	25	10	US-09-866-108-4431	Sequence 4431, Ap
	10	12.6	63.0	25	10	US-09-866-108-4432	Sequence 4432, Ap
	11	12.6	63.0	25	10	US-09-866-108-4433	Sequence 4433, Ap
	12	12.6	63.0	25	10	US-09-866-108-4434	Sequence 4434, Ap
	13	12.6	63.0	25	10	US-09-866-108-4435	Sequence 4435, Ap
	14	12.6	63.0	43	9	US-09-978-295A-450	Sequence 450, App
	15	12.6	63.0	43	9	US-09-978-697-450	Sequence 450, App
	16	12.6	63.0	75	10	US-09-864-761-28813	Sequence 28813, A
	17	12.6	63.0	91	10	US-09-864-761-19573	Sequence 19573, A
c	18	12.4	62.0	18	10	US-09-886-607-17	Sequence 17, Appl
	19	12.4	62.0	36	10	US-09-728-466-26	Sequence 26, Appl
c	20	12.4	62.0	36	10	US-09-728-466-28	Sequence 28, Appl
	21	12.4	62.0	97	10	US-09-923-876-3313	Sequence 3313, Ap
	22	12.4	62.0	99	10	US-09-294-093B-458	Sequence 458, App
	23	12.2	61.0	17	10	US-09-866-108-1500	Sequence 1500, Ap
c	24	12.2	61.0	17	10	US-09-866-108-6279	Sequence 6279, Ap
	25	12.2	61.0	25	10	US-09-866-108-4436	Sequence 4436, Ap
	26	12.2	61.0	25	10	US-09-866-108-4437	Sequence 4437, Ap
c	27	12.2	61.0	25	10	US-09-866-108-11171	Sequence 11171, A
c	28	12.2	61.0	25	10	US-09-866-108-11172	Sequence 11172, A
c	29	12.2	61.0	25	10	US-09-866-108-11173	Sequence 11173, A
c	30	12.2	61.0	25	10	US-09-866-108-11174	Sequence 11174, A
c	31	12.2	61.0	25	10	US-09-866-108-11175	Sequence 11175, A
c	32	12.2	61.0	25	10	US-09-866-108-11176	Sequence 11176, A
c	33	12.2	61.0	25	10	US-09-866-108-11177	Sequence 11177, A
c	34	12.2	61.0	25	10	US-09-866-108-11178	Sequence 11178, A
c	35	12.2	61.0	25	10	US-09-866-108-11179	Sequence 11179, A
	36	12.2	61.0	47	9	US-09-966-277-43	Sequence 43, Appl
c	37	12.2	61.0	47	9	US-09-966-277-44	Sequence 44, Appl
c	38	12.2	61.0	61	9	US-09-946-807-1301	Sequence 1301, Ap
c	39	12.2	61.0	61	10	US-09-795-668-1301	Sequence 1301, Ap
c	40	12.2	61.0	61	10	US-09-795-686-1301	Sequence 1301, Ap
	41	12	60.0	28	10	US-09-852-000-12	Sequence 12, Appl
c	42	12	60.0	28	10	US-09-852-000-13	Sequence 13, Appl
c	43	12	60.0	38	10	US-09-972-331-3	Sequence 3, Appli
c	44	12	60.0	38	10	US-09-972-331-7	Sequence 7, Appli
c	45	12	60.0	47	9	US-09-853-526-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-09-974-300-4020
; Sequence 4020, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4020
; LENGTH: 99
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(99)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-4020

Query Match 77.0%; Score 15.4; DB 10; Length 99;
Best Local Similarity 94.1%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       4 GCACGGCGCAGAGGAGC 20
           ||||| ||||||||||
Db      31 GCACGCCGCAGAGGAGC 47
```

RESULT 2
US-09-894-633A-18
; Sequence 18, Application US/09894633A
; Patent No. US20020124285A1
; GENERAL INFORMATION:

```
;  APPLICANT: Conner, Timothy
;  APPLICANT:  Dubois, Patrice
;  APPLICANT:  Malven, Marianne
;  APPLICANT:  Masucci, James
;  TITLE OF INVENTION: PLANT REGULATORY SEQUENCES FOR SELECTIVE CONTROL OF GENE EXPRESSI
;  FILE REFERENCE: 38-21(15856)B
;  CURRENT APPLICATION NUMBER: US/09/894,633A
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: 60/214,357
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 09/894,633
;  PRIOR FILING DATE: 2000-06-28
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(27)
;   OTHER INFORMATION: synthetic primer sequence
US-09-894-633A-18

  Query Match             76.0%;  Score 15.2;  DB 10;  Length 27;
  Best Local Similarity   85.0%;  Pred. No. 1.6e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           |||| |||||| |||||| |
Db       7 CGAGGACGGCGAAGAGGATC 26


RESULT 3
US-09-886-607-14/c
; Sequence 14, Application US/09886607
; Patent No. US20020083482A1
; GENERAL INFORMATION:
;  APPLICANT: Miller, Marcia
;  APPLICANT:  Afanassieff, Marielle
;  APPLICANT:  Briles, W. Elwood
;  TITLE OF INVENTION: Method for Breeding Disease Resistant Domesticated Fowl
;  FILE REFERENCE: 1954-376
;  CURRENT APPLICATION NUMBER: US/09/886,607
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: US 09/244,093
;  PRIOR FILING DATE: 1999-02-04
;  PRIOR APPLICATION NUMBER: US 09/619,757
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: US 08/774,025
;  PRIOR FILING DATE: 1996-12-27
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Gallus sp.
US-09-886-607-14

  Query Match             72.0%;  Score 14.4;  DB 10;  Length 21;
  Best Local Similarity   93.8%;  Pred. No. 3.7e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CACGGCGCAGAGGAGC 20
           ||||| ||||||||||
Db      16 CACGGTGCAGAGGAGC 1


RESULT 4
US-10-158-895-7
; Sequence 7, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
;  APPLICANT: ONO, KOICHIRO
;  APPLICANT:  OHTOMO, TOSHIHIKO
;  APPLICANT:  TSUCHIYA, MASAYUKI
;  TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
;  FILE REFERENCE: 053466/0278
;  CURRENT APPLICATION NUMBER: US/10/158,895
;  CURRENT FILING DATE:  2002-06-03
;  PRIOR APPLICATION NUMBER: US/09/529,279
;  PRIOR FILING DATE: 2000-04-11
;  PRIOR APPLICATION NUMBER: PCT/JP98/04796
;  PRIOR FILING DATE: 1998-10-22
;  PRIOR APPLICATION NUMBER: JP 9/290188
;  PRIOR FILING DATE: 1997-10-22
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-10-158-895-7

  Query Match             65.0%;  Score 13;  DB 9;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+03;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 CGGCGCAGAGGAG 19
           |||||||||||||
Db      14 CGGCGCAGAGGAG 26


RESULT 5
US-09-972-809-4
; Sequence 4, Application US/09972809
; Patent No. US20020151490A1
; GENERAL INFORMATION:
;  APPLICANT: Sundeep, Khosla
;  APPLICANT:  Conover, Cheryl A.
;  TITLE OF INVENTION: TREATMENT OF OSTEOPOROSIS
;  FILE REFERENCE: 07039/183001
;  CURRENT APPLICATION NUMBER: US/09/972,809
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/428,226
;  PRIOR FILING DATE: 1999-10-27
;  PRIOR APPLICATION NUMBER: 60/045,607
;  PRIOR FILING DATE: 1997-05-05
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 75
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically derived primer
US-09-972-809-4

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 75;
  Best Local Similarity   87.5%;  Pred. No. 1.9e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAG 16
           ||||||||  ||||||
Db      32 CGAGCACGCAGCAGAG 47


RESULT 6
US-09-864-761-19378
; Sequence 19378, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
```

```
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19378
;   LENGTH: 88
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC011297.1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.99
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.9
;   OTHER INFORMATION: EST_HUMAN HIT: BF115285.1, EVALUE 3.00e-16
;   OTHER INFORMATION: NT HIT: J03061.1, EVALUE 2.00e-16
;   OTHER INFORMATION: SWISSPROT HIT: Q90663, EVALUE 9.30e+00
US-09-864-761-19378

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 88;
  Best Local Similarity   87.5%;  Pred. No. 1.9e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 AGCACGGCGCAGAGGA 18
           || | ||||||||||
Db      22 AGGAAGGCGCAGAGGA 37
```

```
RESULT 7
US-09-866-108-4429
; Sequence 4429, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 4429
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-4429

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GAGCACGGCGCAGAGGAGC 20
           | ||| |  ||||||||||
Db       7 GGGCAGGAAGCAGAGGAGC 25


RESULT 8
US-09-866-108-4430
; Sequence 4430, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
```

```
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 4430
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-4430

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GAGCACGGCGCAGAGGAGC 20
           | ||| |  ||||||||||
Db       6 GGGCAGGAAGCAGAGGAGC 24


RESULT 9
US-09-866-108-4431
; Sequence 4431, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 4431
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-4431

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GAGCACGGCGCAGAGGAGC 20
           | ||| |  ||||||||||
Db       5 GGGCAGGAAGCAGAGGAGC 23


RESULT 10
US-09-866-108-4432
; Sequence 4432, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 4432
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-4432

Query Match 63.0%; Score 12.6; DB 10; Length 25;
Best Local Similarity 78.9%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GAGCACGGCGCAGAGGAGC 20
| ||| | ||||||||||
Db 4 GGGCAGGAAGCAGAGGAGC 22

RESULT 11
US-09-866-108-4433
; Sequence 4433, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 4433
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-4433

Query Match 63.0%; Score 12.6; DB 10; Length 25;
Best Local Similarity 78.9%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GAGCACGGCGCAGAGGAGC 20
| ||| | ||||||||||
Db 3 GGGCAGGAAGCAGAGGAGC 21

RESULT 12
US-09-866-108-4434
; Sequence 4434, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine

```
; SEQ ID NO 4434
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-4434

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GAGCACGGCGCAGAGGAGC 20
          | ||| |  ||||||||||
Db      2 GGGCAGGAAGCAGAGGAGC 20


RESULT 13
US-09-866-108-4435
; Sequence 4435, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 4435
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-4435

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GAGCACGGCGCAGAGGAGC 20
          | ||| |  ||||||||||
Db      1 GGGCAGGAAGCAGAGGAGC 19


RESULT 14
US-09-978-295A-450
; Sequence 450, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
;  APPLICANT:  Ferrara, Napoleon
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, J. Christopher
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Kuo, Sophia S.
;  APPLICANT:  Napier, Mary A.
;  APPLICANT:  Pan, James;
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2630P1C11
;  CURRENT APPLICATION NUMBER: US/09/978,295A
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064249
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/077450
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: 60/077632
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077641
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077649
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078004
;  PRIOR FILING DATE: 1998-03-13
;  PRIOR APPLICATION NUMBER: 60/078886
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078936
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078939
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/079294
```

; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 63.0%; Score 12.6; DB 9; Length 43;
Best Local Similarity 78.9%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      2 GAGCACGGCGCAGAGGAGC 20
          ||||||| | || | |||
Db      8 GAGCACGGGGAAGGGCAGC 26
```

RESULT 15
US-09-978-697-450

; Sequence 450, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C27
; CURRENT APPLICATION NUMBER: US/09/978,697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499

; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 63.0%; Score 12.6; DB 9; Length 43;
Best Local Similarity 78.9%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       2 GAGCACGGCGCAGAGGAGC 20
           ||||||||  |  ||  |  |||
Db       8 GAGCACGGGGAAGGGCAGC 26
```

Search completed: November 23, 2002, 06:42:01
Job time : 18.25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-2
Perfect score: 20
Sequence: 1 cgagcacggcgcagaggagc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	18.4	92.0	94	9	AA197619	AA197619 mu19a03.r
	2	15.8	79.0	52	10	AW693858	AW693858 NF069H05S
c	3	15.2	76.0	55	13	BI250613	BI250613 602993676
c	4	15.2	76.0	90	13	BI104703	BI104703 602891614
c	5	14.8	74.0	96	17	GGA270142	AJ231913 Gallus ga
c	6	14.2	71.0	55	17	BH643795	BH643795 1008060C0
c	7	14.2	71.0	79	14	BQ167594	BQ167594 WHE0077_H
c	8	13.8	69.0	92	17	AZ595307	AZ595307 1M0407P17
c	9	13.8	69.0	93	17	CNS039QQ	AL234251 Tetraodon
	10	13.8	69.0	95	17	BH173450	BH173450 0104C7 ma
	11	13.8	69.0	98	13	BI768381	BI768381 603053690
	12	13.6	68.0	46	9	AA008276	AA008276 mg69b08.r
c	13	13.6	68.0	50	9	AU104090	AU104090 AU104090
c	14	13.6	68.0	81	13	BM490089	BM490089 pgp2n.pk0
c	15	13.6	68.0	81	17	BH849516	BH849516 SALK_0698
c	16	13.6	68.0	86	17	BH796584	BH796584 1008094H0
c	17	13.6	68.0	87	17	AZ918012	AZ918012 1006002H0
	18	13.6	68.0	93	14	BQ564305	BQ564305 gi14c02.y
	19	13.4	67.0	25	17	AZ374695	AZ374695 1M0127I08
	20	13.4	67.0	74	9	AA492659	AA492659 vi77a03.r
	21	13.4	67.0	88	9	AA574321	AA574321 nf45f03.s
c	22	13.2	66.0	42	17	AZ633442	AZ633442 1M0488N08
	23	13.2	66.0	75	17	CNS01XAQ	AL171467 Tetraodon
	24	13.2	66.0	79	9	AI000040	AI000040 os46c07.s
c	25	13.2	66.0	85	9	AA128241	AA128241 zl92f09.r
c	26	13.2	66.0	86	12	BF528126	BF528126 602042896
c	27	13.2	66.0	95	17	BH216187	BH216187 1006041A0
c	28	13.2	66.0	100	10	AW998118	AW998118 RC0-BN005
c	29	12.8	64.0	54	17	AZ584073	AZ584073 1M0388D22
c	30	12.8	64.0	63	14	BQ761521	BQ761521 EBro08_SQ
	31	12.8	64.0	88	9	AA883676	AA883676 am12f05.s
c	32	12.8	64.0	88	14	D38686	D38686 HUMCI254 Hu
c	33	12.8	64.0	93	9	AI400715	AI400715 tg93h01.x
	34	12.8	64.0	94	17	AZ921327	AZ921327 1006029D0
c	35	12.8	64.0	96	14	R59502	R59502 yh17f12.r1
	36	12.8	64.0	100	17	AZ819462	AZ819462 2M0091A10
c	37	12.6	63.0	33	17	AZ486766	AZ486766 1M0315H10
c	38	12.6	63.0	47	13	BI082347	BI082347 602877544
	39	12.6	63.0	47	17	BH811174	BH811174 SALK_0575
c	40	12.6	63.0	50	9	AU106272	AU106272 AU106272
c	41	12.6	63.0	50	9	AU106275	AU106275 AU106275
c	42	12.6	63.0	50	9	AU106276	AU106276 AU106276
c	43	12.6	63.0	50	9	AU106279	AU106279 AU106279
c	44	12.6	63.0	50	9	AU106287	AU106287 AU106287
c	45	12.6	63.0	50	9	AA250719	AA250719 mx79b08.r

ALIGNMENTS

RESULT 1
AA197619
LOCUS AA197619 94 bp mRNA linear EST 19-FEB-1997
DEFINITION mu19a03.r1 Soares_thymus_2NbMT Mus musculus cDNA clone IMAGE:639820 5', mRNA sequence.
ACCESSION AA197619
VERSION AA197619.1 GI:1793270
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 94)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

```
            MGI:391812
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 81.
FEATURES              Location/Qualifiers
     source           1. .94
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:639820"
                      /clone_lib="Soares_thymus_2NbMT"
                      /sex="male"
                      /tissue_type="Thymus"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                      3']; double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of the modified pT7T3 vector. RNA
                      provided by Dr. Bertrand Jordan. Library went through two
                      rounds of normalization, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."
BASE COUNT       12 a     32 c     34 g     16 t
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 9;  Length 94;
  Best Local Similarity   95.0%;  Pred. No. 1e+03;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           |||||||||||| |||||||
Db       3 CGAGCACGGCGCGGAGGAGC 22


RESULT 2
AW693858
LOCUS       AW693858                 52 bp    mRNA    linear   EST 21-DEC-2000
DEFINITION  NF069H05ST1F1047 Developing stem Medicago truncatula cDNA clone
            NF069H05ST 5', mRNA sequence.
ACCESSION   AW693858
VERSION     AW693858.2  GI:11958019
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
            ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
            ,R.A.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula stem library
  JOURNAL   Unpublished (2000)
COMMENT     On Apr 14, 2000 this sequence version replaced gi:7568595.
            Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Insert Length: 685   Std Error: 0.00
            Plate: 069  row: H  column: 05
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES              Location/Qualifiers
     source           1. .52
                      /organism="Medicago truncatula"
                      /db_xref="taxon:3880"
                      /clone="NF069H05ST"
                      /clone_lib="Developing stem"
                      /tissue_type="stem"
                      /dev_stage="Pooled developmental"
                      /note="Vector: Lambda Zap; Contains a mixture of
                      internodal stem segments"
BASE COUNT       14 a     18 c     14 g      6 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 10;  Length 52;
  Best Local Similarity   89.5%;  Pred. No. 9.4e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAG 19
           ||||||| ||| |||||||
Db      21 CGAGCACCGCGAAGAGGAG 39


RESULT 3
BI250613/c
LOCUS       BI250613                 55 bp    mRNA    linear   EST 17-JUL-2001
DEFINITION  602993676F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5149483 5',
            mRNA sequence.
ACCESSION   BI250613
VERSION     BI250613.1  GI:14799130
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11369  row: g  column: 20
            High quality sequence start: 2
            High quality sequence stop: 55.
FEATURES              Location/Qualifiers
     source           1. .55
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:5149483"
                      /clone_lib="NCI_CGAP_Mam5"
                      /tissue_type="tumor, gross tissue"
                      /dev_stage="7 months"
                      /lab_host="DH10B"
                      /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Library constructed by Life Technologies. Investigators
                      providing samples: Lothar Hennighausen/Robin Humphreys,
                      NIH"
BASE COUNT        8 a     19 c     17 g     11 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 13;  Length 55;
  Best Local Similarity   85.0%;  Pred. No. 1.6e+04;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           || |||||| |||||||| |
Db      41 CGGGCACGGAGCAGAGGACC 22


RESULT 4
```

```
BI104703/c
LOCUS       BI104703                 90 bp    mRNA    linear   EST 26-JUN-2001
DEFINITION  602891614F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5036690 5',
            mRNA sequence.
ACCESSION   BI104703
VERSION     BI104703.1  GI:14555596
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11101  row: 1  column: 03
            High quality sequence stop: 90.
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Mus musculus"
                     /strain="CZECH II"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5036690"
                     /clone_lib="NCI_CGAP_Lu29"
                     /tissue_type="spontaneous tumor, metastatic to mammary.
                     Stem cell origin."
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT       16 a     29 c     29 g     16 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 13;  Length 90;
  Best Local Similarity   85.0%;  Pred. No. 1.8e+04;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CGAGCACGGCGCAGAGGAGC 20
           || |||||| |||||||| |
Db      75 CGGGCACGGAGCAGAGGACC 56


RESULT 5
GGA270142/c
LOCUS       GGA270142                96 bp    DNA     linear   GSS 08-JUL-2000
DEFINITION  Gallus gallus anonymous sequence from Cosmid mapping to chromosome
            5 (Cosmid 27 - Contig 43), genomic survey sequence.
ACCESSION   AJ231913
VERSION     AJ231913.1  GI:3451731
KEYWORDS    GSS; genome survey sequence.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Smith,J., Bruley,C.K., Paton,I.R., Dunn,I., Jones,C.T., Windsor,D.,
            Morrice,D.R., Law,A.S., Masabanda,J., Sazanov,A., Waddington,D.,
            Fries,R. and Burt,D.W.
  TITLE     Differences in gene density on chicken macrochromosomes and
            microchromosomes
  JOURNAL   Anim. Genet. 31 (2), 96-103 (2000)
  MEDLINE   20244064
    PUBMED  10782207
REFERENCE   2  (bases 1 to 96)
  AUTHORS   Smith,J., Bruley,C.K., Paton,I.R., Law,A.S., Masabanda,J.,
            Waddington,D., Fries,R. and Burt,D.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-AUG-1998) Division of Molecular Biology, Roslin
            Institute, Roslin, Midlothian EH25 9PS, Scotland, UK
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /chromosome="5"
                     /map="5q"
BASE COUNT       12 a     37 c     27 g     20 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 17;  Length 96;
  Best Local Similarity   88.9%;  Pred. No. 2.6e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 AGCACGGCGCAGAGGAGC 20
           ||||||| ||||||| ||
Db      36 AGCACGGAGCAGAGGGGC 19


RESULT 6
BH643795/c
LOCUS       BH643795                 55 bp    DNA     linear   GSS 14-FEB-2002
DEFINITION  1008060C01.2EL_y1 1008 - RescueMu Grid I Zea mays genomic, DNA
            sequence.
ACCESSION   BH643795
VERSION     BH643795.1  GI:18672592
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1008060  row: 21
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1008 - RescueMu Grid I"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the  web
                     site www.zmdb.iastate.edu and follow the links for
                     'RescueMu.' Grid I was grown at Berkeley in 2001. DNA was
                     extracted from leaf punches,  double digested using BamHI
                     and BglII, and ligated to form circular plasmids.  DH10B
                     cells were transformed and then screened on LB plates with
```

ampicillin."
BASE COUNT 3 a 30 c 12 g 10 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 17; Length 55;
Best Local Similarity 84.2%; Pred. No. 4e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GAGCACGGCGCAGAGGAGC 20
||||||| |||| |||||
Db 35 GAGCACGCAGCAGCGGAGC 17

RESULT 7
BQ167594/c
LOCUS BQ167594 79 bp mRNA linear EST 25-APR-2002
DEFINITION WHE0077_H03_P05ZK Wheat endosperm cDNA library Triticum aestivum cDNA clone WHE0077_H03_P05, mRNA sequence.
ACCESSION BQ167594
VERSION BQ167594.1 GI:20312113
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 79)
AUTHORS Altenbach,S., Anderson,O.D., Chao,S., Galili,G., Han,P.S., Hsia ,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.
TITLE The structure and function of the expressed portion of the wheat genomes - Endosperm cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov
This EST was generated by sequencing from the 3' end of the clone. Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20.
Seq primer: Stratagene's KS primer.
FEATURES Location/Qualifiers
source 1. .79
/organism="Triticum aestivum"
/cultivar="Cheyenne"
/db_xref="taxon:4565"
/clone="WHE0077_H03_P05"
/clone_lib="Wheat endosperm cDNA library"
/tissue_type="Endosperm"
/dev_stage="5 to 30 days post anthesis seed"
/lab_host="E. coli SOLR"
/note="Vector: Lambda ZAP II, excised phagemid; Site_1: EcoRI; Seeds collected, endosperm isolated, and RNA prepared by Susan Altenbach. Library constructed by Stratagene, Inc. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab."
BASE COUNT 11 a 31 c 23 g 14 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 79;
Best Local Similarity 84.2%; Pred. No. 4.2e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GAGCACGGCGCAGAGGAGC 20
|||| ||||| ||| ||||
Db 58 GAGCGCGGCGGAGACGAGC 40

RESULT 8
AZ595307/c
LOCUS AZ595307 92 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0407P17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0407P17 R, DNA sequence.
ACCESSION AZ595307
VERSION AZ595307.1 GI:11717497
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 92)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0407 row: P column: 17
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 92.
FEATURES Location/Qualifiers
source 1. .92
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0407P17"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 15 a 21 c 26 g 30 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 92;
Best Local Similarity 88.2%; Pred. No. 6.2e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AGCACGGCGCAGAGGAG 19
| ||||| |||||||||
Db 88 AACACGGGGCAGAGGAG 72

RESULT 9
CNS039QQ/c

```
LOCUS       CNS039QQ                 93 bp    DNA     linear   GSS 15-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            008G11 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL234251
VERSION     AL234251.1  GI:7893386
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 93)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 93)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="008G11"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG008AD06SP1~end :
                     PUC-Ori"
BASE COUNT        0 a     29 c     31 g     31 t      2 others
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 93;
  Best Local Similarity   88.2%;  Pred. No. 6.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GCACGGCGCAGAGGAGC 20
          || || ||||||||||||
Db     20 GCCCGACGCAGAGGAGC 4


RESULT 10
BH173450
LOCUS       BH173450                 95 bp    DNA     linear   GSS 09-OCT-2001
DEFINITION  0104C7 maize genomic ORF library F Zea mays genomic, DNA sequence.
ACCESSION   BH173450
VERSION     BH173450.1  GI:15988814
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Fu,Y., Liu,F. and Schnable,F.S.
  TITLE     Maize genomic sequences rescued by a novel expression vector system
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Patrick S. Schnable
            Iowa State University
            G405 Agronomy, Ames, IA 50011, USA
            Tel: 515 294 0975
            Fax: 515 294 2299
            Email: schnable@iastate.edu
            Size-selected Sau3AI-digested DNA (100-800 bp) from immature,
            unpollinated ears was cloned into the BamHI site of pHis6. Ligation
            products were digested with McrBC prior to transformation.
            Class: plasmid ends.
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
                     /clone_lib="maize genomic ORF library F"
                     /note="Vector: pHis6; Site_1: PstI; Site_2: BamHI"
BASE COUNT       21 a     29 c     28 g     17 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 95;
  Best Local Similarity   88.2%;  Pred. No. 6.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GCACGGCGCAGAGGAGC 20
          | ||||||||||||| ||
Db      5 GTACGGCGCAGAGGCGC 21


RESULT 11
BI768381
LOCUS       BI768381                 98 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603053690F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5203151 5',
            mRNA sequence.
ACCESSION   BI768381
VERSION     BI768381.1  GI:15759946
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11509  row: c  column: 24
            High quality sequence stop: 98.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5203151"
                     /clone_lib="NIH_MGC_122"
                     /lab_host="DH10B"
                     /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                     anonymous pool of 24 week female lung, 16 week female
                     spleen, and 20-22 week male spleens.  Library is oligo-dT
                     primed and directionally cloned (EcoRV site is destroyed
                     upon  cloning).  Average insert size 1.4 kb, insert size
                     range 1-3 kb. Library is normalized and enriched for
                     full-length clones and was  constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking  code 026. Note:
                     this is a NIH_MGC Library."
BASE COUNT       26 a     32 c     32 g      8 t
ORIGIN
```

```
  Query Match             69.0%;  Score 13.8;  DB 13;  Length 98;
  Best Local Similarity   88.2%;  Pred. No. 6.3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGG 17
          || |||||||||| |||
Db     62 CGGGCACGGCGCATAGG 78


RESULT 12
AA008276
LOCUS       AA008276                46 bp    mRNA    linear   EST 25-JUL-1996
DEFINITION  mg69b08.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:438231 5' similar to SW:CRB3_RAT P02524 BETA CRYSTALLIN
            B3. ;, mRNA sequence.
ACCESSION   AA008276
VERSION     AA008276.1  GI:1464247
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:263567
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: ETPrimer
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:438231"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo.  "
BASE COUNT       12 a     13 c     17 g      4 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 46;
  Best Local Similarity   80.0%;  Pred. No. 6.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          |||||||||  |||   ||||
Db      4 CGAGCACGGAGCACCTGAGC 23


RESULT 13
AU104090/c
LOCUS       AU104090                50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU104090 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP07261, mRNA sequence.
ACCESSION   AU104090
VERSION     AU104090.1  GI:13553611
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP07261"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        8 a     17 c     16 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 6.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          || | ||||||| ||| |||
Db     44 CGGGTACGGCGCTGAGCAGC 25


RESULT 14
BM490089/c
LOCUS       BM490089                81 bp    mRNA    linear   EST 07-FEB-2002
DEFINITION  pgp2n.pk002.g22 Normalized Chicken Pituitary/Hypothalamus/Pineal
            Library (pgp2n) Gallus gallus cDNA clone pgp2n.pk002.g22 5' similar
            to no significant hits (pLog(P) 4), mRNA sequence.
ACCESSION   BM490089
VERSION     BM490089.1  GI:18611030
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Porter,T.E. and Cogburn,L.A.
  TITLE     ESTs from Normalized Chicken Pituitary/Hypothalamus/Pineal cDNA
            library, USDA/IFAFS Animal Genome Project
  JOURNAL   Unpublished (2002)
```

```
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .81
                     /organism="Gallus gallus"
                     /strain="Commercial broiler chickens"
                     /db_xref="taxon:9031"
                     /clone="pgp2n.pk002.g22"
                     /clone_lib="Normalized Chicken
                     Pituitary/Hypothalamus/Pineal Library (pgp2n)"
                     /sex="Male and Female"
                     /tissue_type="Pituitary Gland/Hypothalamus/Pineal Gland"
                     /dev_stage="Embryonic (d12,d14,d19); post-hatch (1,3,5,7,9
                     weeks)"
                     /lab_host="E. coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
                     pools of total RNA isolated from each tissue at different
                     ages. Single pass sequencing from 5'-end"
BASE COUNT       12 a     37 c     18 g     11 t      3 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 13;  Length 81;
  Best Local Similarity   80.0%;  Pred. No. 7.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          |||||  |||| || |||||
Db     25 CGAGCGAGGCGGAGCGGAGC 6


RESULT 15
BH849516/c
LOCUS       BH849516                 81 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  SALK_069806.18.35.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_069806.18.35.x, DNA
            sequence.
ACCESSION   BH849516
VERSION     BH849516.1  GI:21420387
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated exon of At3g25080.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .81
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_069806.18.35.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       21 a     20 c     16 g     24 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 81;
  Best Local Similarity   80.0%;  Pred. No. 7.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGAGCACGGCGCAGAGGAGC 20
          ||| || ||| ||||| |||
Db     30 CGACCAAGGCCCAGAGTAGC 11


Search completed: November 26, 2002, 04:06:27
Job time : 768.8 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-3
Perfect score: 20
Sequence: 1 ggacgagggcgagcacggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	16.8	84.0	51	6	AX156995	AX156995 Sequence
c	2	15.2	76.0	99	6	A23183	A23183 Artificial
c	3	14.2	71.0	26	6	AX146654	AX146654 Sequence
c	4	14.2	71.0	40	6	AX490799	AX490799 Sequence
	5	14.2	71.0	45	6	AX128401	AX128401 Sequence
	6	14.2	71.0	45	6	AX429834	AX429834 Sequence
	7	14.2	71.0	65	6	AR038667	AR038667 Sequence
c	8	14.2	71.0	66	6	AR038665	AR038665 Sequence
	9	14.2	71.0	78	14	AF362846	AF362846 HIV-1 iso
	10	14.2	71.0	78	14	AF362847	AF362847 HIV-1 iso
	11	14.2	71.0	78	14	AF362848	AF362848 HIV-1 iso
	12	13.8	69.0	60	6	AR009383	AR009383 Sequence
c	13	13.8	69.0	78	6	A67678	A67678 Sequence 8
	14	13.6	68.0	30	6	AR063475	AR063475 Sequence
	15	13.6	68.0	30	6	I28155	I28155 Sequence 12
	16	13.6	68.0	37	6	AX356235	AX356235 Sequence
	17	13.6	68.0	49	11	AL823803	AL823803 Arabidops
c	18	13.6	68.0	64	9	HUMMACU	L37707 Homo sapien
c	19	13.6	68.0	86	6	AX236733	AX236733 Sequence
c	20	13.6	68.0	90	6	AX236627	AX236627 Sequence
c	21	13.6	68.0	100	11	AF235063	AF235063 Mus muscu
	22	13.2	66.0	21	6	A86893	A86893 Sequence 3
	23	13.2	66.0	21	6	A93725	A93725 Sequence 1
	24	13.2	66.0	21	6	AR088407	AR088407 Sequence
	25	13.2	66.0	27	6	AX207516	AX207516 Sequence
	26	13.2	66.0	27	6	AX207554	AX207554 Sequence
	27	13.2	66.0	60	6	AR009382	AR009382 Sequence
	28	13.2	66.0	75	14	AF178663	AF178663 HIV-1 LCV
	29	13.2	66.0	81	14	AF362871	AF362871 HIV-1 iso
	30	13.2	66.0	81	14	AF362873	AF362873 HIV-1 iso
	31	13.2	66.0	81	14	AF362874	AF362874 HIV-1 iso
c	32	13	65.0	37	6	AR170305	AR170305 Sequence
c	33	13	65.0	37	6	AX007133	AX007133 Sequence
c	34	13	65.0	37	6	E51079	E51079 Novel Esche
c	35	12.6	63.0	22	6	AX328831	AX328831 Sequence
c	36	12.6	63.0	23	6	AR106710	AR106710 Sequence
c	37	12.6	63.0	23	6	AR202521	AR202521 Sequence
c	38	12.6	63.0	23	6	I24358	I24358 Sequence 8
c	39	12.6	63.0	24	6	BD004778	BD004778 Novel VEG
c	40	12.6	63.0	24	6	E36097	E36097 Novel polyp
c	41	12.6	63.0	24	6	E38429	E38429 Novel polyp
c	42	12.6	63.0	25	6	AX083192	AX083192 Sequence
c	43	12.6	63.0	25	6	AX107065	AX107065 Sequence
c	44	12.6	63.0	25	6	AX155273	AX155273 Sequence
c	45	12.6	63.0	25	6	AX206869	AX206869 Sequence

ALIGNMENTS

RESULT 1
AX156995
LOCUS AX156995 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 323 from Patent WO0140521.
ACCESSION AX156995
VERSION AX156995.1 GI:14538326
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 51)
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof

```
  JOURNAL   Patent: WO 0140521-A 323 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (324 is other entry)
                     Accession number cg20436198"
BASE COUNT        5 a     14 c     29 g      3 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 6;  Length 51;
  Best Local Similarity   90.0%;  Pred. No. 2.8e+04;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGCG 20
          ||||| ||||||| ||||||
Db     22 GGACGCGGGCGAGGACGGCG 41


RESULT 2
A23183/c
LOCUS       A23183                   99 bp    DNA     linear   PAT 11-JAN-1995
DEFINITION  Artificial DNA for DNA adapter (id. 7).
ACCESSION   A23183
VERSION     A23183.1  GI:833251
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Stern,A., Hagemann,I. and Ziegler-Landesberger,D.
  TITLE     Method of introduction of nucleic acids in cells
  JOURNAL   Patent: EP 0544292-A 7 02-JUN-1993;
            BOEHRINGER MANNHEIM GMBH
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       13 a     36 c     27 g     23 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 6;  Length 99;
  Best Local Similarity   85.0%;  Pred. No. 9.5e+04;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGCG 20
          |||||| |||||  ||||||
Db     88 GGACGACGGCGACGACGGCG 69


RESULT 3
AX146654/c
LOCUS       AX146654                 26 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 5 from Patent WO0134818.
ACCESSION   AX146654
VERSION     AX146654.1  GI:14285047
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Johal,G.S. and Multani,D.S.
  TITLE     Sorghum dwarfing genes and methods of use
  JOURNAL   Patent: WO 0134818-A 5 17-MAY-2001;
            Johal, Gurmukh S. (US) ; Multani, Dilbag S. (US)
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide primer designed from sequence of
                     Zea mays Br2 gene"
BASE COUNT        0 a     12 c      6 g      8 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 26;
  Best Local Similarity   84.2%;  Pred. No. 2.8e+05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          |||||| | || |||||||
Db     24 GACGAGCGGGAACACGGCG 6


RESULT 4
AX490799/c
LOCUS       AX490799                 40 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 13 from Patent WO0246400.
ACCESSION   AX490799
VERSION     AX490799.1  GI:22323749
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Kranz,D.M., Starwalt,S. and Bluestone,J.A.
  TITLE     Mutated class ii major histocompatibility proteins
  JOURNAL   Patent: WO 0246400-A 13 13-JUN-2002;
            The Board of Trustees of the University of Illinois (US)
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic nucleotide primer"
BASE COUNT        4 a     19 c     12 g      5 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 40;
  Best Local Similarity   84.2%;  Pred. No. 2.6e+05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          |||| ||||||| || |||
Db     39 GACGTGGGCGAGTACCGCG 21


RESULT 5
AX128401
LOCUS       AX128401                 45 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 62 from Patent WO0130843.
ACCESSION   AX128401
VERSION     AX128401.1  GI:14134909
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Barbas,C.F., Kadan,M. and Beerli,R.
  TITLE     Ligand activated transcriptional regulator proteins
  JOURNAL   Patent: WO 0130843-A 62 03-MAY-2001;
            Novartis AG (CH) ; The Scripps Research Institute (US)
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Recombinant molecule"
BASE COUNT        9 a     14 c     17 g      5 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 45;
  Best Local Similarity   84.2%;  Pred. No. 2.5e+05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGC 19
          ||| ||||||| || ||||
```

```
Db      3 GGAGGAGGGCGCGCCCGGC 21


RESULT 6
AX429834
LOCUS       AX429834                 45 bp    DNA     linear   PAT 21-JUN-2002
DEFINITION  Sequence 26 from Patent WO0206463.
ACCESSION   AX429834
VERSION     AX429834.1  GI:21541010
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Beerli,R., Schopfer,U. and Barbas,C.F.
  TITLE     Regulation of gene expression using single-chain, monomeric, ligand
            dependent polypeptide switches
  JOURNAL   Patent: WO 0206463-A 26 24-JAN-2002;
            The Scripps Research Institute (US)
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="unidentified"
                     /db_xref="taxon:32644"
                     /note="Synthesized"
BASE COUNT        9 a     14 c     17 g      5 t
ORIGIN

  Query Match           71.0%;  Score 14.2;  DB 6;  Length 45;
  Best Local Similarity   84.2%;  Pred. No. 2.5e+05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGC 19
          ||| ||||||| || ||||
Db      3 GGAGGAGGGCGCGCCCGGC 21


RESULT 7
AR038667
LOCUS       AR038667                 65 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 5 from patent US 5807675.
ACCESSION   AR038667
VERSION     AR038667.1  GI:5958030
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Davalian,D., Singh,R. and Ullman,E.F.
  TITLE     Fluorescent oxygen channeling immunoassays
  JOURNAL   Patent: US 5807675-A 5 15-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="unknown"
BASE COUNT       16 a     15 c     23 g     11 t
ORIGIN

  Query Match           71.0%;  Score 14.2;  DB 6;  Length 65;
  Best Local Similarity   84.2%;  Pred. No. 2.4e+05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          || | ||||||||| ||||
Db      3 GAAGCGGGCGAGCATGGCG 21


RESULT 8
AR038665/c
LOCUS       AR038665                 66 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 3 from patent US 5807675.
ACCESSION   AR038665
VERSION     AR038665.1  GI:5958028
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Davalian,D., Singh,R. and Ullman,E.F.
  TITLE     Fluorescent oxygen channeling immunoassays
  JOURNAL   Patent: US 5807675-A 3 15-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="unknown"
BASE COUNT       14 a     26 c      8 g     18 t
ORIGIN

  Query Match           71.0%;  Score 14.2;  DB 6;  Length 66;
  Best Local Similarity   84.2%;  Pred. No. 2.4e+05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          || | ||||||||| ||||
Db     28 GAAGCGGGCGAGCATGGCG 10


RESULT 9
AF362846
LOCUS       AF362846                 78 bp    RNA     linear   VRL 09-MAY-2001
DEFINITION  HIV-1 isolate 3040P1 from USA envelope glycoprotein, V3 region
            (env) gene, partial cds.
ACCESSION   AF362846
VERSION     AF362846.1  GI:14010375
KEYWORDS    .
SOURCE      Human immunodeficiency virus type 1.
  ORGANISM  Human immunodeficiency virus type 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Freel,S.A., Williams,J.M., Nelson,J.A., Patton,L.L., Fiscus,S.A.,
            Swanstrom,R. and Shugars,D.C.
  TITLE     Characterization of human immunodeficiency virus type 1 in saliva
            and blood plasma by V3-specific heteroduplex tracking assay and
            genotype analyses
  JOURNAL   J. Virol. 75 (10), 4936-4940 (2001)
  MEDLINE   21211648
   PUBMED   11312368
REFERENCE   2  (bases 1 to 78)
  AUTHORS   Freel,S.A., Williams,J.M., Nelson,J.A., Patton,L.L., Fiscus,S.A.,
            Swanstrom,R. and Shugars,D.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2001) Dentistry and Microbiology, University of
            North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Human immunodeficiency virus type 1"
                     /virion
                     /isolate="3040P1"
                     /db_xref="taxon:11676"
                     /tissue_type="blood plasma"
                     /country="USA"
     gene            <1. .>78
                     /gene="env"
     CDS             <1. .>78
                     /gene="env"
                     /note="V3 region"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAK51973.1"
                     /db_xref="GI:14010376"
                     /translation="NNNTRRSIQIGPGRAWHTTRIIGDIR"
BASE COUNT       39 a     10 c     18 g     11 t
ORIGIN

  Query Match           71.0%;  Score 14.2;  DB 14;  Length 78;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy      1 GGACGAGGGCGAGCACGGC 19
          |||| |||| ||||| |||
Db     31 GGACCAGGGAGAGCATGGC 49


RESULT 10
AF362847
LOCUS       AF362847                 78 bp    RNA     linear   VRL 09-MAY-2001
DEFINITION  HIV-1 isolate 3040P2 from USA envelope glycoprotein, V3 region
            (env) gene, partial cds.
ACCESSION   AF362847
VERSION     AF362847.1  GI:14010377
KEYWORDS    .
SOURCE      Human immunodeficiency virus type 1.
  ORGANISM  Human immunodeficiency virus type 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Freel,S.A., Williams,J.M., Nelson,J.A., Patton,L.L., Fiscus,S.A.,
            Swanstrom,R. and Shugars,D.C.
  TITLE     Characterization of human immunodeficiency virus type 1 in saliva
            and blood plasma by V3-specific heteroduplex tracking assay and
            genotype analyses
  JOURNAL   J. Virol. 75 (10), 4936-4940 (2001)
  MEDLINE   21211648
   PUBMED   11312368
REFERENCE   2  (bases 1 to 78)
  AUTHORS   Freel,S.A., Williams,J.M., Nelson,J.A., Patton,L.L., Fiscus,S.A.,
            Swanstrom,R. and Shugars,D.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2001) Dentistry and Microbiology, University of
            North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Human immunodeficiency virus type 1"
                     /virion
                     /isolate="3040P2"
                     /db_xref="taxon:11676"
                     /tissue_type="blood plasma"
                     /country="USA"
     gene            <1. .>78
                     /gene="env"
     CDS             <1. .>78
                     /gene="env"
                     /note="V3 region"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAK51974.1"
                     /db_xref="GI:14010378"
                     /translation="NNNTRRSIQIGPGRAWHTTRIIGDIR"
BASE COUNT       39 a     11 c     18 g     10 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 14;  Length 78;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGC 19
          |||| |||| ||||| |||
Db     31 GGACCAGGGAGAGCATGGC 49


RESULT 11
AF362848
LOCUS       AF362848                 78 bp    RNA     linear   VRL 09-MAY-2001
DEFINITION  HIV-1 isolate 3040P3 from USA envelope glycoprotein, V3 region
            (env) gene, partial cds.
ACCESSION   AF362848
VERSION     AF362848.1  GI:14010379
KEYWORDS    .
SOURCE      Human immunodeficiency virus type 1.
  ORGANISM  Human immunodeficiency virus type 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Freel,S.A., Williams,J.M., Nelson,J.A., Patton,L.L., Fiscus,S.A.,
            Swanstrom,R. and Shugars,D.C.
  TITLE     Characterization of human immunodeficiency virus type 1 in saliva
            and blood plasma by V3-specific heteroduplex tracking assay and
            genotype analyses
  JOURNAL   J. Virol. 75 (10), 4936-4940 (2001)
  MEDLINE   21211648
   PUBMED   11312368
REFERENCE   2  (bases 1 to 78)
  AUTHORS   Freel,S.A., Williams,J.M., Nelson,J.A., Patton,L.L., Fiscus,S.A.,
            Swanstrom,R. and Shugars,D.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2001) Dentistry and Microbiology, University of
            North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Human immunodeficiency virus type 1"
                     /virion
                     /isolate="3040P3"
                     /db_xref="taxon:11676"
                     /tissue_type="blood plasma"
                     /country="USA"
     gene            <1. .>78
                     /gene="env"
     CDS             <1. .>78
                     /gene="env"
                     /note="V3 region"
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAK51975.1"
                     /db_xref="GI:14010380"
                     /translation="NNNTRRSIQIGPGRAWHTTRTIGDIR"
BASE COUNT       39 a     11 c     18 g     10 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 14;  Length 78;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGC 19
          |||| |||| ||||| |||
Db     31 GGACCAGGGAGAGCATGGC 49


RESULT 12
AR009383
LOCUS       AR009383                 60 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 151 from patent US 5756291.
ACCESSION   AR009383
VERSION     AR009383.1  GI:3968188
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Griffin,L., Albrecht,G., Latham,J., Leung,L., Vermaas,E. and
            Toole,J.J.
  TITLE     Aptamers specific for biomolecules and methods of making
  JOURNAL   Patent: US 5756291-A 151 26-MAY-1998;
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="unknown"
BASE COUNT       13 a      6 c     32 g      9 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 60;
  Best Local Similarity   88.2%;  Pred. No. 3.4e+05;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACG 17
          ||| |||||||||| |||
```

```
Db      28 GGAAGAGGGCGAGGACG 44


RESULT 13
A67678/c
LOCUS       A67678                    78 bp    DNA     linear   PAT 05-MAY-1999
DEFINITION  Sequence 8 from Patent WO9743428.
ACCESSION   A67678
VERSION     A67678.1  GI:4756541
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Merot,B., Baudino,S., Gruber,V., Bournat,P., Lenee,P., Riviere,M.E.
            and Audonnet,J.F.
  TITLE     TRANSGENIC PLANTS EXPRESSING RABIES GLYCOPROTEIN G, AND
            GLYCOPROTEINS THUS OBTAINED
  JOURNAL   Patent: WO 9743428-A 8 20-NOV-1997;
            BIOCEM (FR)
COMMENT     Other publication FR 2748480 19971114.
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       10 a     32 c     26 g     10 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 78;
  Best Local Similarity   88.2%;  Pred. No. 3.2e+05;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CGAGGGCGAGCACGGCG 20
           ||| |||||| ||||||
Db      58 CGAAGGCGAGGACGGCG 42


RESULT 14
AR063475
LOCUS       AR063475                  30 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 12 from patent US 5846706.
ACCESSION   AR063475
VERSION     AR063475.1  GI:5992783
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Greenberg,S.J. and Evans,M.Jo.
  TITLE     Rapid and sensitive detection of herpes simplex virus
  JOURNAL   Patent: US 5846706-A 12 08-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        7 a      6 c     17 g      0 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 4.5e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGCG 20
           ||||| || |||| |||| |
Db       1 GGACGGGGACGAGGACGGGG 20


RESULT 15
I28155
LOCUS       I28155                    30 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 12 from patent US 5569583.
ACCESSION   I28155
VERSION     I28155.1  GI:1818931
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Greenberg,S.J. and Evans,M.J.
  TITLE     Rapid and sensitive detection of cytomegalovirus
  JOURNAL   Patent: US 5569583-A 12 29-OCT-1996;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        7 a      6 c     17 g      0 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 4.5e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGCG 20
           ||||| || |||| |||| |
Db       1 GGACGGGGACGAGGACGGGG 20


Search completed: December  3, 2002, 18:12:09
Job time : 309.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-3
Perfect score: 20
Sequence: 1 ggacgagggcgagcacggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31433	Human neuropilin m
	2	16.8	84.0	51	22	AAI73382	Human silent SNP c
c	3	15.2	76.0	95	14	AAQ42957	Encodes 8 Arg C-te
c	4	14.2	71.0	26	22	AAD06342	Sorghum dwarfing g
c	5	14.2	71.0	30	21	AAZ98642	Probe for E. coli
c	6	14.2	71.0	30	21	AAZ98666	Escherichia coli K
c	7	14.2	71.0	40	24	ABN84043	Single chain Class
	8	14.2	71.0	45	22	AAD06100	Drosophila EcR cod
	9	14.2	71.0	45	24	AAI72470	D. melanogaster ec
	10	14.2	71.0	50	21	AAZ98641	E. coli K12 DNAJ o
	11	14.2	71.0	50	21	AAZ98665	Escherichia coli K
c	12	14.2	71.0	60	23	AAS48592	Pseudomonas aerugi
c	13	14.2	71.0	60	23	AAS48615	Pseudomonas aerugi
c	14	14.2	71.0	66	16	AAQ85876	"CTAATC-30mer" mod
	15	13.8	69.0	60	19	AAV60517	Cloned Factor X-bi
c	16	13.8	69.0	78	19	AAV19397	Signal sequence fr
	17	13.6	68.0	30	17	AAT47994	Herpes simplex vir
	18	13.6	68.0	30	20	AAV72560	Herpes simplex vir
	19	13.6	68.0	37	24	ABK27836	Corn male reproduc
	20	13.6	68.0	44	15	AAQ64210	Degenerate probe f
c	21	13.6	68.0	51	22	AAL31268	Human SNP oligonuc
c	22	13.6	68.0	51	22	AAL31269	Human SNP oligonuc
c	23	13.6	68.0	86	22	AAS44166	Neisseria meningit
c	24	13.6	68.0	90	22	AAS44060	Neisseria meningit
	25	13.4	67.0	49	21	AAA30428	Human ACAM cDNA PC
	26	13.4	67.0	50	20	AAX52327	Probe used to isol
	27	13.4	67.0	50	22	AAF72485	Human PRO polypept
	28	13.4	67.0	50	22	AAC97450	Human PRO258 hybri
c	29	13.4	67.0	60	18	AAT76132	Human eosinophil p
c	30	13.4	67.0	60	20	AAX53942	Intercellular adhe
c	31	13.4	67.0	60	21	AAF19504	Human eosinophil p
c	32	13.4	67.0	60	21	AAA33382	Low adenosine anti
	33	13.4	67.0	60	24	ABN34868	Human spliced tran
	34	13.2	66.0	21	18	AAT73118	Primer acbD3 for A
	35	13.2	66.0	21	19	AAV56645	Actinoplanes sp. a
	36	13.2	66.0	27	22	AAH76446	Maize root transcr
	37	13.2	66.0	27	22	AAH26310	Maize root transcr
	38	13.2	66.0	31	24	ABK69149	Human phosphorylas
c	39	13.2	66.0	34	22	AAI68141	Lepomis centrarchi
	40	13.2	66.0	60	19	AAV60516	Cloned Factor X-bi
c	41	13.2	66.0	84	21	AAC30338	Human secreted pro
c	42	13	65.0	37	21	AAZ57918	Escherichia coli t
c	43	12.8	64.0	24	22	AAH78622	PCR primer RFC680(
c	44	12.8	64.0	26	14	AAQ52208	Neuroblastoma spec
	45	12.8	64.0	40	21	AAZ95979	Polynucleotide seq

ALIGNMENTS

RESULT 1
AAZ31433
ID AAZ31433 standard; DNA; 20 BP.
XX
AC AAZ31433;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3603.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 4 A; 5 C; 11 G; 0 U; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGCG 20
           ||||||||||||||||||||
Db       1 GGACGAGGGCGAGCACGGCG 20


RESULT 2
AAI73382
ID   AAI73382 standard; DNA; 51 BP.
XX
AC   AAI73382;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:323.
XX
KW   Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW   protein therapy; vaccine; probe; diagnostic assay; detection;
KW   quantitation; restorative therapy; polymorphic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200140521-A2.
XX
PD   07-JUN-2001.
XX
PF   30-NOV-2000; 2000WO-US32758.
XX
PR   30-NOV-1999;   99US-0168138.
PR   29-NOV-2000; 2000US-0726173.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-356160/37.
XX
PT   Polymorphic nucleic acid sequences, useful in genetic testing and
PT   therapy -
XX
PS   Claim 1; Page 153; 2653pp; English.
XX
CC   AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC   sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC   AAM53114 to AAM53329 represent peptides related to human polymorphic
CC   polynucleotide sequences. The sequences can be used in gene and protein
CC   therapy, and in vaccine production. (I) and the polypeptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of polymorphic polypeptides.
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 51 BP; 5 A; 14 C; 29 G; 3 T; 0 other;

  Query Match             84.0%;  Score 16.8;  DB 22;  Length 51;
  Best Local Similarity   90.0%;  Pred. No. 4.9e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGCG 20
           ||||| ||||||| ||||||
Db      22 GGACGCGGGCGAGGACGGCG 41


RESULT 3
AAQ42957/c
ID   AAQ42957 standard; DNA; 95 BP.
XX
AC   AAQ42957;
XX
DT   07-SEP-1993  (first entry)
XX
DE   Encodes 8 Arg C-terminal extension of G-CSF.
XX
KW   Granulocyte Colony Stimulating Factor; Cell-homing factor;
KW   fusion protein; targetting; gene therapy; ds.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    63..86
FT                   /*tag=  a
FT                   /note= "8 Arg codons -
FT                          there is a 4-nucleotide overhang at the
FT                          5' end of the coding strand and a 4-nt
FT                          overhang (5'-AGCT-3') at the 5' end of the
FT                          complementary strand"
XX
PN   EP544292-A.
XX
PD   02-JUN-1993.
XX
PF   26-NOV-1992;   92EP-0120205.
XX
PR   27-NOV-1991;   91DE-4139001.
XX
PA   (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX
PI   Hagemann I,  Stern A,  Ziegler-Landesberger D;
XX
DR   WPI; 1993-177078/22.
XX
PT   Cell introduction of cell homing factor-polycation-nucleic acid
PT   complex - of linear genetic cell homing factor-polycation fusion
PT   having ionic interaction with nucleic acid, the polycation part
PT   having at least three lysine and arginine
XX
PS   Example 2; Page 6; 17pp; German.
XX
CC   This sequence codes for the C-terminal region of G-CSF fused to 8
CC   Arginine residues and a Proline residue. The basic nature of
CC   the extension ("polycation chain") results in ionic interaction of
CC   the fusion protein with DNA to form a complex. The G-CSF acts to
CC   target the complex to e.g. NFS60 cells where the associated DNA can
CC   be taken up by those cells at a much higher rate than when the basic
```

```
CC   extension is not present.
XX
SQ   Sequence 95 BP; 12 A; 35 C; 26 G; 22 T; 0 other;

  Query Match             76.0%;  Score 15.2;  DB 14;  Length 95;
  Best Local Similarity   85.0%;  Pred. No. 2.1e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGCG 20
          |||||| |||||  ||||||
Db     88 GGACGACGGCGACGACGGCG 69


RESULT 4
AAD06342/c
ID   AAD06342 standard; DNA; 26 BP.
XX
AC   AAD06342;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Sorghum dwarfing gene, Dw3 allele amplifying PCR primer #1.
XX
KW   Sorghum; dwarfing gene; Dw3; P-glycoprotein; transgenic plant;
KW   plant growth; pathogen resistance; plant breeding; PCR primer; ss.
XX
OS   Sorghum bicolor.
XX
PN   WO200134818-A2.
XX
PD   17-MAY-2001.
XX
PF   10-NOV-2000; 2000WO-US30816.
XX
PR   12-NOV-1999;   99US-0165176.
XX
PA   (JOHA/) JOHAL G S.
PA   (MULT/) MULTANI D S.
XX
PI   Johal GS,  Multani DS;
XX
DR   WPI; 2001-329088/34.
XX
PT   Novel sorghum dwarfing genes useful for modifying the growth of the
PT   organisms, particularly plants, such as maize, wheat, rice, Basmati
PT   rice, rye, millet and barley and for enhancing resistance to pathogens
PT   -
XX
PS   Example 1; Page 63; 76pp; English.
XX
CC   The present sequence is a PCR primer used for amplifying sorghum
CC   dwarfing gene, Dw3 allele. The primer is derived from the nucleotide
CC   sequence of maize Br2 gene. Dw3 DNA is useful for modifying the growth
CC   of an organism, especially reducing the height of monocotyledon plants,
CC   such as maize, wheat, rice, Basmati rice, sorghum, rye, millet and
CC   barley. Dw3 DNA is useful for producing transgenic monocotyledon and
CC   dicotyledon plants like soybean, sunflower, safflower, alfalfa, cotton,
CC   Brassica sp., peanuts or fruit trees. The Dw3 DNA is also useful for
CC   isolating nucleotide homologue molecules that encode P-glycoproteins,
CC   enhancing the resistance of plants to pathogens, including bacteria,
CC   fungi, viruses, nematodes and insects, as probes for the isolation of
CC   other P-glycoprotein-like genes and as molecular markers. Identifying
CC   plants that possess a mutant allele is useful in agriculture,
CC   particularly in breeding dwarf crop plants, particularly dwarf sorghum
CC   plants.
XX
SQ   Sequence 26 BP; 0 A; 12 C; 6 G; 8 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 26;
  Best Local Similarity   84.2%;  Pred. No. 5.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          ||||||  | ||  |||||||
Db     24 GACGAGCGGGAACACGGCG 6


RESULT 5
AAZ98642/c
ID   AAZ98642 standard; DNA; 30 BP.
XX
AC   AAZ98642;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Probe for E. coli DNAJ gene fragment target sequence.
XX
KW   Homogeneous specific binding assay; DNAJ gene; analyte determination;
KW   antigen; hapten; drug; cancer marker; pesticide; pollutant;
KW   photochemically activatable chemiluminescent compound; probe; ss.
XX
OS   Escherichia coli.
XX
PN   EP984281-A2.
XX
PD   08-MAR-2000.
XX
PF   21-MAY-1992;   99EP-0121547.
XX
PR   22-MAY-1991;   91US-0704569.
PR   20-JUN-1991;   91US-0718490.
PR   21-MAY-1992;   92EP-0304630.
XX
PA   (DADE-) DADE BEHRING MARBURG GMBH.
XX
PI   Ullman EF,  Kirakossian H,  Pease JS,  Wagner DB,  Daniloff Y;
XX
DR   WPI; 2000-197307/18.
XX
PT   Homogeneous specific binding assay, e.g. for proteins or nucleic acids,
PT   uses photochemically activatable chemiluminescent compound  -
XX
PS   Example 5; Page 66; 78pp; English.
XX
CC   This sequence represents a probe for a fragment of the Escherichia coli
CC   K12 DNAJ gene, used as an example of a target sequence in a homogeneous
CC   assay. The invention relates to a homogeneous assay for determining an
CC   analyte, comprising combining medium suspected of containing an analyte
CC   with a label reagent, and intrinsically metastable species, and
CC   determining the reaction. The label reagent comprises a suspended
CC   particle and a specific binding pair (sbp) member associated with a
CC   photochemically activatable chemiluminescent compound (PACC), the sbp is
CC   capable of binding to a second sbp or to the analyte, the second sbp
CC   being capable of binding to the analyte. The metastable species is
CC   capable of diffusing into the medium and reacts preferentially with the
CC   PACC, when brought into close proximity by the presence or absence of the
CC   analyte. The method can be used to determine antigens or haptens, e.g.
CC   blood group or HLA antigens or bacterial, fungal, protozoal or viral
CC   antigens, other proteins e.g. immunoglobulins, cytokines, enzymes,
CC   hormones, cancer markers or nutritional markers, microorganisms, drugs,
CC   metabolites, pesticides, pollutants or polynucleotides. The method is a
CC   homogeneous assay and does not require a separation step.
XX
SQ   Sequence 30 BP; 2 A; 14 C; 8 G; 6 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 21;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 5.8e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          || | |||||||||  ||||
Db     28 GAAGCGGGCGAGCATGGCG 10


RESULT 6
AAZ98666/c
```

```
ID   AAZ98666 standard; DNA; 30 BP.
XX
AC   AAZ98666;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Escherichia coli K12 DNAJ gene fragment probe sequence.
XX
KW   Specific binding assay; DNAJ; analyte determination; antigen;
KW   hapten; drug; cancer marker; pesticide; chemiluminescent compound;
KW   pollutant; probe; ss.
XX
OS   Escherichia coli.
XX
PN   EP984282-A2.
XX
PD   08-MAR-2000.
XX
PF   21-MAY-1992;   99EP-0121551.
XX
PR   22-MAY-1991;   91US-0704569.
PR   20-JUN-1991;   91US-0718490.
PR   21-MAY-1992;   92EP-0304630.
XX
PA   (DADE-) DADE BEHRING MARBURG GMBH.
XX
PI   Ullman EF,  Kirakossian H,  Pease JS,  Wagner DB,  Daniloff Y;
XX
DR   WPI; 2000-225922/20.
XX
PT   Particles containing a chemiluminescent donor and a fluorescent
PT   acceptor are useful in specific binding assays to determine antigens,
PT   haptens, enzymes, hormones, cancer markers or nutritional markers -
XX
PS   Example 5; Page 66; 79pp; English.
XX
CC   This sequence represents a probe for a fragment of the Escherichia coli
CC   DNAJ gene. The fragment is used as an example target sequence in a method
CC   for determining analyte in a sample using the particles of the invention.
CC   The invention relates to particles (I) containing a compound (II) and a
CC   fluorescent compound (III). Compound II reacts with singlet oxygen to
CC   form a metastable intermediate that can decompose with simultaneous or
CC   subsequent emission of light. The fluorescent compound (III) is excited
CC   by activated compound (II) and emits at a wavelength longer than the
CC   emission wavelength of compound (II). The particles are useful in
CC   specific binding assays. The assays can be used to determine antigens or
CC   haptens, e.g. blood group or HLA antigens or bacterial, fungal, protozoal
CC   or viral antigens, other proteins e.g. immunoglobulins, cytokines,
CC   enzymes, hormones, cancer markers or nutritional markers, microorganisms,
CC   drugs, metabolites, pesticides, pollutants or polynucleotides. The longer
CC   emission wavelength of (III) eliminates interference from serum
CC   components when the particles are used in specific binding assays in the
CC   presence of such components.
XX
SQ   Sequence 30 BP; 2 A; 14 C; 8 G; 6 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 21;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 5.8e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           || | ||||||||| ||||
Db      28 GAAGCGGGCGAGCATGGCG 10


RESULT 7
ABN84043/c
ID   ABN84043 standard; DNA; 40 BP.
XX
AC   ABN84043;
XX
DT   23-SEP-2002  (first entry)
XX
DE   Single chain Class II MHC, I-Ag7 mutagenic primer 4.
XX
KW   Major histocompatibility complex; MHC; combinatorial library;
KW   I-Ag7; insulin-dependent diabetes mellitus; autoimmune disease;
KW   antidiabetic; diagnosis; mutagenesis; PCR; primer; ss.
XX
OS   Synthetic.
XX
PN   WO200246400-A2.
XX
PD   13-JUN-2002.
XX
PF   10-DEC-2001; 2001WO-US47830.
XX
PR   08-DEC-2000; 2000US-254248P.
XX
PA   (UNII ) UNIV ILLINOIS FOUND.
XX
PI   Kranz DM,  Starwalt S,  Bluestone JA;
XX
DR   WPI; 2002-537566/57.
XX
PT   Novel isolated mutant major histocompatibility complex class II
PT   chimeric protein displayed on recombinant yeast cell surface has
PT   improved stability or T-cell receptor binding, useful for treating
PT   autoimmune disease  -
XX
PS   Example 3; Page 27; 66pp; English.
XX
CC   The present sequence is mutagenic primer 3 for positions beta-56
CC   and beta-57 of a single chain (sc) Class II major histocompatibility
CC   complex (MHC) haplotype I-Ag7 construct.  A site-directed mutagenesis
CC   strategy involving PCR sewing and yeast homologous recombination was
CC   used to mutate residues beta-56 and -57.  Mutant libraries of sc MHC
CC   Class II protein were generated and displayed on the cell surface of
CC   recombinant yeast cells, and stabilised I-Ag7 proteins were
CC   identified.  Haplotype I-Ag7 is associated with the development of
CC   insulin-dependent diabetes mellitus (IDDM) in the non-obese (NOD)
CC   mouse model of IDDM.  The present invention allows the creation and
CC   isolation of stabilised variants of Class II peptide-MHC complexes.
CC   Stabilised variants of single chain I-Ag7, in association with each
CC   of 3 peptides of interest (BDC2.5(A)), GAD65(78-95) and B9-23
CC   insulin), have been produced.  These can be used in diagnostic
CC   tests of IDDM, in imaging tests, and in the treatment or prevention
CC   of IDDM.
XX
SQ   Sequence 40 BP; 4 A; 19 C; 12 G; 5 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 24;  Length 40;
  Best Local Similarity   84.2%;  Pred. No. 5.7e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           |||| ||||||| || |||
Db      39 GACGTGGGCGAGTACCGCG 21


RESULT 8
AAD06100
ID   AAD06100 standard; DNA; 45 BP.
XX
AC   AAD06100;
XX
DT   31-JUL-2001  (first entry)
XX
DE   Drosophila EcR coding region amplifying primer, (AscI)-EcR.
XX
KW   Fusion protein; nucleotide-binding domain; NBD;
KW   ligand-binding domain; LBD; transcription regulating domain; TRD;
KW   zinc finger protein; ZFP; ligand-activated transcriptional regulator;
KW   gene regulation; gene therapy; cell proliferative disorder; cancer;
KW   psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW   EcR; PCR primer; ecdysone receptor; ss.
```

XX
OS Drosophila sp.
XX
PN WO200130843-A1.
XX
PD 03-MAY-2001.
XX
PF 23-OCT-2000; 2000WO-EP10430.
XX
PR 25-OCT-1999; 99US-0433042.
PR 02-JUN-2000; 2000US-0586625.
XX
PA (NOVS) NOVARTIS AG.
PA (SCRI) SCRIPPS RES INST.
XX
PI Barbas CF, Kadan M, Beerli R;
XX
DR WPI; 2001-308618/32.
XX
PT New fusion protein containing nucleotide-binding and ligand-binding
PT domains, useful e.g. in gene therapy of cancer, provides
PT ligand-activated control of gene expression -
XX
PS Example 12; Page 110; 218pp; English.
XX
CC The invention relates to fusion protein comprising a nucleotide-binding
CC domain (NBD), a ligand-binding domain (LBD) of an intracellular
CC receptor (ICR) and a transcription regulating domain (TRD). NBD is a
CC polydactyl zinc finger protein, or a modular part of it, that interacts
CC specifically with a contiguous sequence of at least 3 nucleotides. The
CC fusion protein functions as a ligand-activated transcriptional regulator.
CC The fusion protein and the nucleic acid encoding it, are used to regulate
CC gene expression, particularly in gene therapy for treating malignant
CC cell proliferative diseases (e.g. colon cancer, prostate cancer,
CC renal-cell carcinoma) and non-malignant cell proliferative
CC diseases (e.g. psoriasis, pemphigus vulgaris, Behcet's syndrome and
CC lipid histiocytosis). The fusion protein and its DNA are also useful for
CC treating diseases caused by viruses in humans/plants, genetic and/or
CC acquired diseases. The fusion protein can be designed to target any
CC selected gene (endogenous or exogenous), and can be made to have
CC different selectivity or specificity for endogenous or exogenous ligands.
CC The present sequence is Drosophila ecdysone receptor (EcR) coding region
CC amplifying primer, (AscI)-EcR. This primer is used to obtain LBD which
CC is used to construct fusion protein of the invention.
XX
SQ Sequence 45 BP; 9 A; 14 C; 17 G; 5 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 45;
Best Local Similarity 84.2%; Pred. No. 5.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGACGAGGGCGAGCACGGC 19
||| ||||||| || ||||
Db 3 GGAGGAGGGCGCGCCCGGC 21

RESULT 9
AAI72470
ID AAI72470 standard; DNA; 45 BP.
XX
AC AAI72470;
XX
DT 16-MAY-2002 (first entry)
XX
DE D. melanogaster ecdysone receptor LBD primer, EcRAsc-B.
XX
KW Polymerase chain reaction; primer; amplify; PCR; reporter;
KW polypeptide gene switch; ligand binding domain; LBD; transcription;
KW nuclear hormone receptor; gene therapy; regulation; expression; ss.
XX
OS Synthetic.
XX
PN WO200206463-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-EP08190.
XX
PR 18-JUL-2000; 2000US-0619063.
XX
PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PA (SCRI) SCRIPPS RES INST.
XX
PI Beerli R, Schopfer U, Barbas CF;
XX
DR WPI; 2002-179793/23.
XX
PT Single chain, monomeric polypeptide gene switches for regulating gene
PT function, comprise two ligand binding domains derived from nuclear
PT hormone receptors operatively linked to one functional domain -
XX
PS Example 1; Page 31; 63pp; English.
XX
CC The sequences given in AAI72463-72 are primers which were used in the
CC generation of single chain switch which could be used in the
CC construction of a non-naturally occurring polypeptide gene switch
CC comprising two ligand binding domains (LBDs) derived from nuclear
CC hormone receptors operatively linked to a first functional domain. The
CC gene switch is useful for regulating the function of a target nucleotide
CC that contains a defined sequence, by exposing the target nucleotide to
CC the polypeptide gene switch in the presence of a ligand that binds one
CC of the LBDs of the polypeptide, where the DNA binding domain of the
CC polypeptide binds the defined sequence, or the functional domain of
CC the polypeptide alters the function of the target nucleotide. The gene
CC switch is also useful in the field of gene therapy and as a regulator
CC of gene expression or transcription. The advantage of the gene switches
CC of the invention over existing gene switches is the need for only a
CC single molecular switch and a single expression vector for production
CC of that switch.
XX
SQ Sequence 45 BP; 9 A; 14 C; 17 G; 5 T; 0 other;

Query Match 71.0%; Score 14.2; DB 24; Length 45;
Best Local Similarity 84.2%; Pred. No. 5.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGACGAGGGCGAGCACGGC 19
||| ||||||| || ||||
Db 3 GGAGGAGGGCGCGCCCGGC 21

RESULT 10
AAZ98641
ID AAZ98641 standard; DNA; 50 BP.
XX
AC AAZ98641;
XX
DT 05-JUN-2000 (first entry)
XX
DE E. coli K12 DNAJ oligonucleotide used as a target in a homogeneous assay.
XX
KW Homogeneous specific binding assay; DNAJ gene; analyte determination;
KW antigen; hapten; drug; cancer marker; pesticide; pollutant;
KW photochemically activatable chemiluminescent compound; ss.
XX
OS Escherichia coli.
XX
PN EP984281-A2.
XX
PD 08-MAR-2000.
XX
PF 21-MAY-1992; 99EP-0121547.
XX
PR 22-MAY-1991; 91US-0704569.
PR 20-JUN-1991; 91US-0718490.

```
PR   21-MAY-1992;   92EP-0304630.
XX
PA   (DADE-) DADE BEHRING MARBURG GMBH.
XX
PI   Ullman EF,  Kirakossian H,  Pease JS,  Wagner DB,  Daniloff Y;
XX
DR   WPI; 2000-197307/18.
XX
PT   Homogeneous specific binding assay, e.g. for proteins or nucleic acids,
PT   uses photochemically activatable chemiluminescent compound  -
XX
PS   Example 5; Page 65; 78pp; English.
XX
CC   This sequence represents a fragment of the Escherichia coli K12 DNAJ
CC   gene, used as an example of a target sequence in a homogeneous assay. The
CC   invention relates to a homogeneous assay for determining an analyte,
CC   comprising combining medium suspected of containing an analyte with a
CC   label reagent, and intrinsically metastable species, and determining the
CC   reaction. The label reagent comprises a suspended particle and a
CC   specific binding pair (sbp) member associated with a photochemically
CC   activatable chemiluminescent compound (PACC), the sbp is capable of
CC   binding to a second sbp or to the analyte, the second sbp being capable
CC   of binding to the analyte. The metastable species is capable of diffusing
CC   into the medium and reacts preferentially with the PACC, when brought
CC   into close proximity by the presence or absence of the analyte. The
CC   method can be used to determine antigens or haptens, e.g. blood group or
CC   HLA antigens or bacterial, fungal, protozoal or viral antigens, other
CC   proteins e.g. immunoglobulins, cytokines, enzymes, hormones, cancer
CC   markers or nutritional markers, microorganisms, drugs, metabolites,
CC   pesticides, pollutants or polynucleotides. The method is a homogeneous
CC   assay and does not require a separation step.
XX
SQ   Sequence 50 BP; 10 A; 12 C; 23 G; 5 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 21;  Length 50;
  Best Local Similarity   84.2%;  Pred. No. 5.6e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           || | ||||||||| ||||
Db      13 GAAGCGGGCGAGCATGGCG 31


RESULT 11
AAZ98665
ID   AAZ98665 standard; DNA; 50 BP.
XX
AC   AAZ98665;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Escherichia coli K12 DNAJ gene fragment nucleotide sequence.
XX
KW   Specific binding assay; DNAJ; analyte determination; antigen;
KW   hapten; drug; cancer marker; pesticide; chemiluminescent compound;
KW   pollutant; ss.
XX
OS   Escherichia coli.
XX
PN   EP984282-A2.
XX
PD   08-MAR-2000.
XX
PF   21-MAY-1992;   99EP-0121551.
XX
PR   22-MAY-1991;   91US-0704569.
PR   20-JUN-1991;   91US-0718490.
PR   21-MAY-1992;   92EP-0304630.
XX
PA   (DADE-) DADE BEHRING MARBURG GMBH.
XX
PI   Ullman EF,  Kirakossian H,  Pease JS,  Wagner DB,  Daniloff Y;
XX
DR   WPI; 2000-225922/20.
XX
PT   Particles containing a chemiluminescent donor and a fluorescent
PT   acceptor are useful in specific binding assays to determine antigens,
PT   haptens, enzymes, hormones, cancer markers or nutritional markers -
XX
PS   Example 5; Page 66; 79pp; English.
XX
CC   This sequence represents a fragment of the Escherichia coli DNAJ gene.
CC   The fragment is used as an example target sequence in a method for
CC   determining analyte in a sample using the particles of the invention. The
CC   invention relates to particles (I) containing a compound (II) and a
CC   fluorescent compound (III). Compound II reacts with singlet oxygen to
CC   form a metastable intermediate that can decompose with simultaneous or
CC   subsequent emission of light. The fluorescent compound (III) is excited
CC   by activated compound (II) and emits at a wavelength longer than the
CC   emission wavelength of compound (II). The particles are useful in
CC   specific binding assays. The assays can be used to determine antigens or
CC   haptens, e.g. blood group or HLA antigens or bacterial, fungal, protozoal
CC   or viral antigens, other proteins e.g. immunoglobulins, cytokines,
CC   enzymes, hormones, cancer markers or nutritional markers, microorganisms,
CC   drugs, metabolites, pesticides, pollutants or polynucleotides. The longer
CC   emission wavelength of (III) eliminates interference from serum
CC   components when the particles are used in specific binding assays in the
CC   presence of such components.
XX
SQ   Sequence 50 BP; 10 A; 12 C; 23 G; 5 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 21;  Length 50;
  Best Local Similarity   84.2%;  Pred. No. 5.6e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           || | ||||||||| ||||
Db      13 GAAGCGGGCGAGCATGGCG 31


RESULT 12
AAS48592/c
ID   AAS48592 standard; DNA; 60 BP.
XX
AC   AAS48592;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Pseudomonas aeruginosa cellular proliferation inhibitory sequence #78.
XX
KW   Antisense; ss; prokaryotic cellular proliferation;
KW   antibiotic; antibacterial; drug design.
XX
OS   Pseudomonas aeruginosa.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
```

```
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 1; Seq ID No 1169; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence is an antisense
CC   oligonucleotide of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 3 A; 22 C; 23 G; 12 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 23;  Length 60;
  Best Local Similarity   84.2%;  Pred. No. 5.5e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGC 19
           |||| ||||| |||||| |
Db      47 GGACCAGGGCCAGCACGCC 29


RESULT 13
AAS48615/c
ID   AAS48615 standard; DNA; 60 BP.
XX
AC   AAS48615;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Pseudomonas aeruginosa cellular proliferation inhibitory sequence #101.
XX
KW   Antisense; ss; prokaryotic cellular proliferation;
KW   antibiotic; antibacterial; drug design.
XX
OS   Pseudomonas aeruginosa.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 1; Seq ID No 1192; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence is an antisense
CC   oligonucleotide of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 3 A; 22 C; 23 G; 12 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 23;  Length 60;
  Best Local Similarity   84.2%;  Pred. No. 5.5e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGC 19
           |||| ||||| |||||| |
Db      47 GGACCAGGGCCAGCACGCC 29


RESULT 14
AAQ85876/c
ID   AAQ85876 standard; DNA; 66 BP.
XX
AC   AAQ85876;
XX
DT   06-NOV-1995  (first entry)
XX
DE   "CTAATC-30mer" modified tailed 30 mer oligonucleotide.
XX
KW   CTAATC-30mer; nucleic acid detection method; probe; ss.
XX
OS   Synthetic.
XX
PN   WO9506877-A.
XX
PD   09-MAR-1995.
XX
PF   29-AUG-1994;   94WO-US09705.
XX
PR   03-SEP-1993;   93US-0117365.
XX
PA   (SYNT ) SYNTEX USA INC.
XX
PI   Davalian D,  Singh R,  Ullman EF;
XX
DR   WPI; 1995-115552/15.
XX
PT   Determn of specific binding analyte using photo-sensitiser to
PT   generate singlet oxygen  - which converts precursor to
PT   photoactive indicator when irradiated, also new precursor cpds,
PT   useful in immunoassays and nucleic acid hybridisation assays
XX
PS   Example 1; Page 50; 72pp; English.
XX
CC   AAQ85874-Q85877 are modified oligonucleotide probes, used to
CC   demonstrate a new nucleic acid detection method.
XX
```

```
SQ   Sequence 66 BP; 14 A; 26 C; 8 G; 18 T; 0 other;

  Query Match            71.0%;  Score 14.2;  DB 16;  Length 66;
  Best Local Similarity  84.2%;  Pred. No. 5.5e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          || | |||||||||  ||||
Db     28 GAAGCGGGCGAGCATGGCG 10


RESULT 15
AAV60517
ID   AAV60517 standard; DNA; 60 BP.
XX
AC   AAV60517;
XX
DT   08-DEC-1998  (first entry)
XX
DE   Cloned Factor X-binding aptamer sequence.
XX
KW   Factor X; aptamer; therapeutic; diagnosis; secondary; ss.
XX
OS   Synthetic.
XX
PN   US5756291-A.
XX
PD   26-MAY-1998.
XX
PF   07-JUN-1995;   95US-0484192.
XX
PR   21-AUG-1992;   92US-0934387.
PR   21-FEB-1992;   92WO-US01383.
PR   07-JUN-1995;   95US-0484192.
XX
PA   (GILE-) GILEAD SCI INC.
XX
PI   Albrecht G,  Griffin L,  Latham J,  Leung L,  Toole JJ;
PI   Vermaas E;
XX
DR   WPI; 1998-321524/28.
XX
PT   Assay for thrombin and purification of thrombin - using DNA aptamer
XX
PS   Example 22; Fig 6; 115pp; English.
XX
CC   AAV60515-47 represent cloned Factor X-binding aptamers. The Factor
CC   X-binding aptamers are identified using the method of the
CC   invention. The specification describes a method for identifying
CC   oligomer sequences which specifically bind target molecules such
CC   as serum proteins, kinins, eicosanoids and extracellular proteins.
CC   The method involves complexation of the target molecule with a
CC   mixture of oligonucleotides containing random sequences and sequences
CC   which serve as primer for PCR amplification. A complex is only formed
CC   with specifically binding oligonucleotide sequences. The complex is
CC   isolated, and complexed members of the oligonucleotide mixture are
CC   recovered by PCR. The method can be used to generate aptamers that can
CC   be used for therapeutic and diagnostic purposes, and for generating
CC   secondary aptamers.
XX
SQ   Sequence 60 BP; 13 A; 6 C; 32 G; 9 T; 0 other;

  Query Match            69.0%;  Score 13.8;  DB 19;  Length 60;
  Best Local Similarity  88.2%;  Pred. No. 8e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACG 17
          ||| |||||||||  |||
Db     28 GGAAGAGGGCGAGGACG 44


Search completed: November 23, 2002, 06:28:34
Job time : 102.6 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-3
Perfect score: 20
Sequence: 1 ggacgagggcgagcacggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	30	1	US-07-718-490-2	Sequence 2, Appli
	2	14.2	71.0	50	1	US-07-718-490-1	Sequence 1, Appli
	3	14.2	71.0	65	1	US-08-479-743-5	Sequence 5, Appli
	4	14.2	71.0	65	5	PCT-US94-09705-5	Sequence 5, Appli
c	5	14.2	71.0	66	1	US-08-479-743-3	Sequence 3, Appli
c	6	14.2	71.0	66	5	PCT-US94-09705-3	Sequence 3, Appli
	7	13.8	69.0	60	1	US-08-484-192-151	Sequence 151, App
	8	13.6	68.0	30	1	US-08-428-370A-12	Sequence 12, Appl
	9	13.6	68.0	30	2	US-08-600-764-12	Sequence 12, Appl
	10	13.2	66.0	21	2	US-08-816-105A-3	Sequence 3, Appli
	11	13.2	66.0	31	4	US-09-657-453A-6	Sequence 6, Appli
	12	13.2	66.0	60	1	US-08-484-192-150	Sequence 150, App
c	13	13	65.0	37	4	US-09-344-888A-19	Sequence 19, Appl
c	14	12.6	63.0	23	1	US-08-064-271-8	Sequence 8, Appli
c	15	12.6	63.0	23	3	US-08-930-589A-8	Sequence 8, Appli
c	16	12.6	63.0	23	4	US-09-599-781-8	Sequence 8, Appli
c	17	12.6	63.0	24	4	US-09-182-450-15	Sequence 15, Appl
c	18	12.6	63.0	25	4	US-09-323-873A-30	Sequence 30, Appl
	19	12.6	63.0	31	1	US-08-623-891-12	Sequence 12, Appl
	20	12.6	63.0	31	4	US-09-340-861-12	Sequence 12, Appl
	21	12.6	63.0	31	4	US-09-634-262-12	Sequence 12, Appl
c	22	12.6	63.0	45	1	US-08-594-808B-3	Sequence 3, Appli
c	23	12.6	63.0	45	3	US-08-718-904-50	Sequence 50, Appl
c	24	12.6	63.0	45	5	PCT-US95-10973A-43	Sequence 43, Appl
	25	12.6	63.0	48	4	US-08-979-608A-36	Sequence 36, Appl
c	26	12.6	63.0	51	1	US-08-594-808B-1	Sequence 1, Appli
c	27	12.6	63.0	67	2	US-08-888-497-23	Sequence 23, Appl
c	28	12.6	63.0	67	4	US-09-362-230-23	Sequence 23, Appl
c	29	12.6	63.0	67	5	PCT-US94-07926-23	Sequence 23, Appl
c	30	12.6	63.0	72	6	5198345-25	Patent No. 5198345
c	31	12.6	63.0	73	1	US-07-982-712-12	Sequence 12, Appl
c	32	12.6	63.0	75	3	US-08-718-904-49	Sequence 49, Appl
c	33	12.6	63.0	75	5	PCT-US95-10973A-42	Sequence 42, Appl
c	34	12.6	63.0	84	2	US-08-374-652C-93	Sequence 93, Appl
	35	12.6	63.0	84	2	US-08-374-652C-94	Sequence 94, Appl
	36	12.6	63.0	84	4	US-08-979-608A-37	Sequence 37, Appl
c	37	12.4	62.0	30	4	US-08-793-701-16	Sequence 16, Appl
	38	12.4	62.0	63	5	PCT-US95-13975-49	Sequence 49, Appl
c	39	12.4	62.0	84	4	US-09-404-390-7	Sequence 7, Appli
c	40	12.2	61.0	18	3	US-09-106-038A-89	Sequence 89, Appl
c	41	12.2	61.0	25	1	US-08-211-430-5	Sequence 5, Appli
c	42	12.2	61.0	26	2	US-08-887-798-17	Sequence 17, Appl
c	43	12.2	61.0	37	1	US-09-640-304-1	Sequence 1, Appli
c	44	12.2	61.0	37	1	US-08-190-103-1	Sequence 1, Appli
c	45	12.2	61.0	37	1	US-08-256-966-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-07-718-490-2/c
; Sequence 2, Application US/07718490
; Patent No. 5340716
; GENERAL INFORMATION:
; APPLICANT: Ullman, Edwin F.
; APPLICANT: Kirakossian, Hrair
; APPLICANT: Pease, John S.
; APPLICANT: Daniloff, Yuri
; APPLICANT: Wagner, Daniel B.
; TITLE OF INVENTION: Assay Method Utilizing Photoactivated
; TITLE OF INVENTION: Chemiluminescent Label
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Syntex (U.S.A.) Inc.
; STREET: 3401 Hillview Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/718,490
; FILING DATE: 19910620
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Leitereg, Theodore J.
; REGISTRATION NUMBER: 28,319
; REFERENCE/DOCKET NUMBER: 27340/ DO-1557
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 852-1091
; TELEFAX: (415) 496-3529
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: K12 DNAJ
; INDIVIDUAL ISOLATE: Synthethic

```
US-07-718-490-2

  Query Match             71.0%;  Score 14.2;  DB 1;  Length 30;
  Best Local Similarity   84.2%;  Pred. No. 1.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           || | ||||||||| ||||
Db      28 GAAGCGGGCGAGCATGGCG 10


RESULT 2
US-07-718-490-1
; Sequence 1, Application US/07718490
; Patent No. 5340716
;  GENERAL INFORMATION:
;    APPLICANT:  Ullman, Edwin F.
;    APPLICANT:  Kirakossian, Hrair
;    APPLICANT:  Pease, John S.
;    APPLICANT:  Daniloff, Yuri
;    APPLICANT:  Wagner, Daniel B.
;    TITLE OF INVENTION:  Assay Method Utilizing Photoactivated
;    TITLE OF INVENTION:  Chemiluminescent Label
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Syntex (U.S.A.) Inc.
;      STREET:  3401 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/718,490
;      FILING DATE:  19910620
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leitereg, Theodore J.
;      REGISTRATION NUMBER:  28,319
;      REFERENCE/DOCKET NUMBER:  27340/ DO-1557
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 852-1091
;      TELEFAX:  (415) 496-3529
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)T
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:
;      ORGANISM:  Escherichia coli
;      STRAIN:  K12 DNAJ
;      INDIVIDUAL ISOLATE:  Synthethic
US-07-718-490-1

  Query Match             71.0%;  Score 14.2;  DB 1;  Length 50;
  Best Local Similarity   84.2%;  Pred. No. 1.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           || | ||||||||| ||||
Db      13 GAAGCGGGCGAGCATGGCG 31


RESULT 3
US-08-479-743-5
; Sequence 5, Application US/08479743
; Patent No. 5807675
;  GENERAL INFORMATION:
;    APPLICANT:  Davalian, Dariush
;    APPLICANT:  Singh, Rajendra
;    APPLICANT:  Ullman, Edwin F.
;    TITLE OF INVENTION:  Fluorescent Oxygen Channeling Assays
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Syntex (U.S.A.) Inc.
;      STREET:  3401 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,743
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/117,365
;      FILING DATE:  03-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Roth, Carol J.
;      REGISTRATION NUMBER:  32,783
;      REFERENCE/DOCKET NUMBER:  27870
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 852-1698
;      TELEFAX:  (415) 496-3529
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  65 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  synthetic
US-08-479-743-5

  Query Match             71.0%;  Score 14.2;  DB 1;  Length 65;
  Best Local Similarity   84.2%;  Pred. No. 1.1e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           || | ||||||||| ||||
Db       3 GAAGCGGGCGAGCATGGCG 21


RESULT 4
PCT-US94-09705-5
; Sequence 5, Application PC/TUS9409705
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  FLUORESCENT OXYGEN CHANNELING IMMUNOASSAYS
;    NUMBER OF SEQUENCES:  5
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/09705
;      FILING DATE:  29-AUG-1994
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  65 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  synthetic
PCT-US94-09705-5

  Query Match             71.0%;  Score 14.2;  DB 5;  Length 65;
  Best Local Similarity   84.2%;  Pred. No. 1.1e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           || | ||||||||| ||||
Db       3 GAAGCGGGCGAGCATGGCG 21


RESULT 5
US-08-479-743-3/c
; Sequence 3, Application US/08479743
; Patent No. 5807675
;  GENERAL INFORMATION:
;    APPLICANT:  Davalian, Dariush
;    APPLICANT:  Singh, Rajendra
;    APPLICANT:  Ullman, Edwin F.
;    TITLE OF INVENTION:  Fluorescent Oxygen Channeling Assays
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Syntex (U.S.A.) Inc.
;      STREET:  3401 Hillview Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,743
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/117,365
;      FILING DATE:  03-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Roth, Carol J.
;      REGISTRATION NUMBER:  32,783
;      REFERENCE/DOCKET NUMBER:  27870
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 852-1698
;      TELEFAX:  (415) 496-3529
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  66 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  synthetic
US-08-479-743-3

  Query Match             71.0%;  Score 14.2;  DB 1;  Length 66;
  Best Local Similarity   84.2%;  Pred. No. 1.1e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           || | ||||||||| ||||
Db      28 GAAGCGGGCGAGCATGGCG 10


RESULT 6
PCT-US94-09705-3/c
; Sequence 3, Application PC/TUS9409705
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  FLUORESCENT OXYGEN CHANNELING IMMUNOASSAYS
;    NUMBER OF SEQUENCES:  5
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/09705
;      FILING DATE:  29-AUG-1994
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  66 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  synthetic
PCT-US94-09705-3

  Query Match             71.0%;  Score 14.2;  DB 5;  Length 66;
  Best Local Similarity   84.2%;  Pred. No. 1.1e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           || | ||||||||| ||||
Db      28 GAAGCGGGCGAGCATGGCG 10


RESULT 7
US-08-484-192-151
; Sequence 151, Application US/08484192
; Patent No. 5756291
;  GENERAL INFORMATION:
;    APPLICANT:  GRIFFIN, LINDA C.
;    APPLICANT:  ALBRECHT, GLENN
;    APPLICANT:  LATHAM, JOHN
;    APPLICANT:  LEUNG, LAWRENCE
;    APPLICANT:  VERMAAS, ERIC
;    APPLICANT:  TOOLE, JOHN J.
;    TITLE OF INVENTION:  APTAMERS SPECIFIC FOR BIOMOLECULES AND
;    TITLE OF INVENTION:  METHODS OF MAKING
;    NUMBER OF SEQUENCES:  181
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,192
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US 07/934,387
;       FILING DATE:  21-AUG-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  GRACEY, NANCY J.
;       REGISTRATION NUMBER:  28,216
;       REFERENCE/DOCKET NUMBER:  246102002221
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-813-5600
;       TELEFAX:  415-494-0792
;       TELEX:  706141
;   INFORMATION FOR SEQ ID NO:  151:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  60 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-484-192-151

  Query Match             69.0%;  Score 13.8;  DB 1;  Length 60;
  Best Local Similarity   88.2%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACG 17
           ||| ||||||||| |||
Db      28 GGAAGAGGGCGAGGACG 44


RESULT 8
US-08-428-370A-12
; Sequence 12, Application US/08428370A
; Patent No. 5569583
;  GENERAL INFORMATION:
;    APPLICANT:  Greenberg, Steven J.
;    APPLICANT:  Evans, Mary Jo
;    TITLE OF INVENTION:  Rapid and Sensitive Detection of
;    TITLE OF INVENTION:  Herpesviruses
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hodgson, Russ, Andrews, Woods & Goodyear
;      STREET:  1800 One M&T Plaza
;      CITY:  Buffalo
;      STATE:  New York
;      COUNTRY:  United States
;      ZIP:  14203-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 Mb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS/ Microsoft Windows
;      SOFTWARE:  Wordperfect for Windows 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/428,370A
;      FILING DATE:  25/04/95
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nelson, M. Bud
;      REGISTRATION NUMBER:  35,300
;      REFERENCE/DOCKET NUMBER:  03551.0015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716)856-4000
;      TELEFAX:  716-849-0349
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    ORIGINAL SOURCE:
;      ORGANISM:  Herpes simplex virus 2
;    FEATURE:
;      LOCATION:  Polymerase gene sequence bases 2268-2297
US-08-428-370A-12

  Query Match             68.0%;  Score 13.6;  DB 1;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGCG 20
           ||||| || |||| |||| |
Db       1 GGACGGGGACGAGGACGGGG 20


RESULT 9
US-08-600-764-12
; Sequence 12, Application US/08600764
; Patent No. 5846706
;  GENERAL INFORMATION:
;    APPLICANT:  Greenberg, Steven J.
;    APPLICANT:  Evans, Mary Jo
;    TITLE OF INVENTION:  Rapid and Sensitive Detection of
;    TITLE OF INVENTION:  Herpesviruses
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hodgson, Russ, Andrews, Woods & Goodyear
;      STREET:  1800 One M&T Plaza
;      CITY:  Buffalo
;      STATE:  New York
;      COUNTRY:  United States
;      ZIP:  14203-2391
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS/ Microsoft Windows
;      SOFTWARE:  Wordperfect for Windows 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/600,764
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/428,370
;      FILING DATE:  25/04/95
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nelson, M. Bud
;      REGISTRATION NUMBER:  35,300
;      REFERENCE/DOCKET NUMBER:  03551.0016
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716)856-4000
;      TELEFAX:  716-849-0349
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single-stranded
;      TOPOLOGY:  linear
;    ORIGINAL SOURCE:
;      ORGANISM:  Herpes simplex virus 2
;    FEATURE:
;      LOCATION:  Polymerase gene sequence bases 2268-2297
US-08-600-764-12

  Query Match             68.0%;  Score 13.6;  DB 2;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGCG 20
           ||||| || |||| |||| |
Db       1 GGACGGGGACGAGGACGGGG 20


RESULT 10
US-08-816-105A-3
; Sequence 3, Application US/08816105A
; Patent No. 5989882
;  GENERAL INFORMATION:
;    APPLICANT:  Crueger, Anneliese; Dellweg, Hans-Georg; Lenz,
;    APPLICANT:  Jurgen; Schroder, Werner; Pape, Hermann;
```

```
;    APPLICANT:  Goeke, Klaus; Schaper, Beate; Hemker, Michael;
;    APPLICANT:  Piepersberg, Wolfgang; Distler, Jurgen;
;    APPLICANT:  Stratmann, Ansgar
;    TITLE OF INVENTION:  PROCESSES FOR PREPARING ACARVIOSYL
;    TITLE OF INVENTION:  TRANSFERASE AND FOR USING IT IN THE
;    TITLE OF INVENTION:  CONVERSION OF ACARBOSE HOMOLOGUES
;    TITLE OF INVENTION:  INTO ACARBOSE, FOR THE PREPARATION
;    TITLE OF INVENTION:  OF ACARBOSE HOMOLOGUES
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sprung Kramer Schaefer & Briscoe
;      STREET:  660 White Plains Road
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10591-5144
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4 Mb storage
;      COMPUTER:  Apple Macintosh 6500
;      OPERATING SYSTEM:  System 7.5
;      SOFTWARE:  WordPerfect 3.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/816,105A
;      FILING DATE:  14-MAR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 19625269.5
;      FILING DATE:  25-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 19611252.4
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kurt G. Briscoe
;      REGISTRATION NUMBER:  33,141
;      REFERENCE/DOCKET NUMBER:  Bayer 9814-KGB
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (914) 332-1700
;      TELEFAX:  (914) 332-1844
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-816-105A-3

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 21;
  Best Local Similarity   83.3%;  Pred. No. 3.1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 ACGAGGGCGAGCACGGCG 20
          || ||| |||| ||||||
Db      1 ACCAGGCCGAGGACGGCG 18


RESULT 11
US-09-657-453A-6
; Sequence 6, Application US/09657453A
; Patent No. 6458591
; GENERAL INFORMATION:
;  APPLICANT:      Brett P. Monia
;  APPLICANT:  Jacqueline Wyatt
;  TITLE OF INVENTION:      ANTISENSE MODULATION OF PHOSPHORYLASE KINASE ALPHA 2 EXPRESS
;  FILE REFERENCE:      RTS-0136
;  CURRENT APPLICATION NUMBER: US/09/657,453A
;  CURRENT FILING DATE:  2000-09-07
;  NUMBER OF SEQ ID NOS:      105
; SEQ ID NO 6
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR Probe
US-09-657-453A-6

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 31;
  Best Local Similarity   83.3%;  Pred. No. 3e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGC 19
          || |||| ||||||| ||
Db      2 GATGAGGTCGAGCACTGC 19


RESULT 12
US-08-484-192-150
; Sequence 150, Application US/08484192
; Patent No. 5756291
;  GENERAL INFORMATION:
;    APPLICANT:  GRIFFIN, LINDA C.
;    APPLICANT:  ALBRECHT, GLENN
;    APPLICANT:  LATHAM, JOHN
;    APPLICANT:  LEUNG, LAWRENCE
;    APPLICANT:  VERMAAS, ERIC
;    APPLICANT:  TOOLE, JOHN J.
;    TITLE OF INVENTION:  APTAMERS SPECIFIC FOR BIOMOLECULES AND
;    TITLE OF INVENTION:  METHODS OF MAKING
;    NUMBER OF SEQUENCES:  181
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,192
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/934,387
;      FILING DATE:  21-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GRACEY, NANCY J.
;      REGISTRATION NUMBER:  28,216
;      REFERENCE/DOCKET NUMBER:  246102002221
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-813-5600
;      TELEFAX:  415-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  150:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  60 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-484-192-150

  Query Match             66.0%;  Score 13.2;  DB 1;  Length 60;
  Best Local Similarity   83.3%;  Pred. No. 2.8e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGG 18
          ||| |||||||| || |
Db     28 GGAAGAGGGCGAGGACCG 45


RESULT 13
US-09-344-888A-19/c
; Sequence 19, Application US/09344888A
; Patent No. 6291245
```

```
; GENERAL INFORMATION:
;  APPLICANT:  Kopetzki, Erhard
;  APPLICANT:  Schantz, Christian
;  TITLE OF INVENTION:  New Host-Vector System
;  FILE REFERENCE:  CD20315
;  CURRENT APPLICATION NUMBER: US/09/344,888A
;  CURRENT FILING DATE:  1999-06-25
;  PRIOR APPLICATION NUMBER:  EP98113156.8
;  PRIOR FILING DATE:  1998-07-15
;  PRIOR APPLICATION NUMBER:  EP98119078.8
;  PRIOR FILING DATE:  1998-10-09
;  NUMBER OF SEQ ID NOS:  24
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 19
;   LENGTH: 37
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-344-888A-19

  Query Match             65.0%;  Score 13;  DB 4;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+03;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       8 GGCGAGCACGGCG 20
           |||||||||||||
Db      30 GGCGAGCACGGCG 18


RESULT 14
US-08-064-271-8/c
; Sequence 8, Application US/08064271
; Patent No. 5543297
;  GENERAL INFORMATION:
;    APPLICANT:  Kennedy, Brian P.
;    APPLICANT:  Cromlish, Wanda A.
;    APPLICANT:  Mancini, Joseph A.
;    APPLICANT:  O'Neil, Gary
;    APPLICANT:  Vickers, Philip J.
;    APPLICANT:  Wong, Elizabeth
;    TITLE OF INVENTION:  HUMAN CYCLOOXYGENASE-2 cDNA AND
;    TITLE OF INVENTION:  ASSAY FOR EVALUATING CYCLOOXYGENASE ACTIVITY
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  126 Lincoln Avenue
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 in, 1.4kb
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  System 7
;      SOFTWARE:  Microsoft Word 5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/064,271
;      FILING DATE:  19930506
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Panzer, Curtis C.
;      REGISTRATION NUMBER:  33,752
;      REFERENCE/DOCKET NUMBER:  18906IA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908)594-3199
;      TELEFAX:  (908)594-4720
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-064-271-8

  Query Match             63.0%;  Score 12.6;  DB 1;  Length 23;
  Best Local Similarity   78.9%;  Pred. No. 5.3e+03;
  Matches  15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGC 19
           || || |||||||||  ||
Db      21 GGGCGCGGGCGAGCATCGC 3


RESULT 15
US-08-930-589A-8/c
; Sequence 8, Application US/08930589A
; Patent No. 6107087
;  GENERAL INFORMATION:
;    APPLICANT:  MERCK FROSST CANADA & CO.
;    APPLICANT:  O'NEILL, GARY P.
;    APPLICANT:  MANCINI, JOSEPH A.
;    TITLE OF INVENTION:  HIGH LEVEL EXPRESSION OF HUMAN
;    TITLE OF INVENTION:  CYCLOOXYGENASE-2
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000, 126 E. Lincoln Ave.
;      CITY:  Rahway
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/930,589A
;      FILING DATE:  28-JUN-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coppola, Joseph A
;      REGISTRATION NUMBER:  38,413
;      REFERENCE/DOCKET NUMBER:  19029PC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-6734
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-930-589A-8

  Query Match             63.0%;  Score 12.6;  DB 3;  Length 23;
  Best Local Similarity   78.9%;  Pred. No. 5.3e+03;
  Matches  15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGC 19
           || || |||||||||  ||
Db      21 GGGCGCGGGCGAGCATCGC 3


Search completed: November 23, 2002, 06:35:59
Job time : 22.55 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-3
Perfect score: 20
Sequence: 1 ggacgagggcgagcacggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	40	9	US-10-015-536-13	Sequence 13, Appl
	2	14.2	71.0	45	9	US-09-908-153B-26	Sequence 26, Appl
c	3	14.2	71.0	60	10	US-09-815-242-1169	Sequence 1169, Ap
c	4	14.2	71.0	60	10	US-09-815-242-1192	Sequence 1192, Ap
	5	13.6	68.0	37	10	US-09-894-633A-29	Sequence 29, Appl
	6	13.4	67.0	50	9	US-09-905-291A-89	Sequence 89, Appl
	7	13.4	67.0	50	10	US-09-909-320-89	Sequence 89, Appl
	8	13.4	67.0	50	10	US-09-909-088B-89	Sequence 89, Appl
	9	13.2	66.0	27	10	US-09-772-656-4	Sequence 4, Appli
c	10	12.8	64.0	96	10	US-09-969-373-679	Sequence 679, App
c	11	12.6	63.0	25	10	US-09-780-053-15	Sequence 15, Appl
c	12	12.6	63.0	25	10	US-09-779-308-15	Sequence 15, Appl
	13	12.6	63.0	34	10	US-09-852-922-6	Sequence 6, Appli
c	14	12.6	63.0	34	10	US-09-852-922-7	Sequence 7, Appli
	15	12.6	63.0	48	10	US-09-962-055-36	Sequence 36, Appl
	16	12.6	63.0	48	12	US-10-023-529-36	Sequence 36, Appl
	17	12.6	63.0	48	12	US-10-023-523-36	Sequence 36, Appl
	18	12.6	63.0	84	10	US-09-962-055-37	Sequence 37, Appl
	19	12.6	63.0	84	12	US-10-023-529-37	Sequence 37, Appl
	20	12.6	63.0	84	12	US-10-023-523-37	Sequence 37, Appl
c	21	12.6	63.0	90	9	US-09-971-980-17	Sequence 17, Appl
	22	12.4	62.0	59	9	US-09-933-797-527	Sequence 527, App
	23	12.2	61.0	31	10	US-09-801-274-1524	Sequence 1524, Ap
c	24	12.2	61.0	46	10	US-09-972-331-6	Sequence 6, Appli
c	25	12.2	61.0	46	10	US-09-972-331-20	Sequence 20, Appl
	26	12.2	61.0	51	9	US-09-908-153B-16	Sequence 16, Appl
c	27	12.2	61.0	88	10	US-09-923-876-1505	Sequence 1505, Ap
	28	12.2	61.0	94	10	US-09-294-093B-4852	Sequence 4852, Ap
c	29	12.2	61.0	98	10	US-09-969-373-1435	Sequence 1435, Ap
	30	12	60.0	18	10	US-09-962-055-35	Sequence 35, Appl
	31	12	60.0	18	12	US-10-023-529-35	Sequence 35, Appl
	32	12	60.0	18	12	US-10-023-523-35	Sequence 35, Appl
	33	12	60.0	59	10	US-09-922-217-976	Sequence 976, App
	34	12	60.0	59	10	US-09-833-263-976	Sequence 976, App
	35	12	60.0	60	10	US-09-962-055-18	Sequence 18, Appl
	36	12	60.0	60	12	US-10-023-529-18	Sequence 18, Appl
	37	12	60.0	60	12	US-10-023-523-18	Sequence 18, Appl
c	38	12	60.0	86	10	US-09-867-701-3669	Sequence 3669, Ap
	39	12	60.0	90	10	US-09-962-055-34	Sequence 34, Appl
	40	12	60.0	90	12	US-10-023-529-34	Sequence 34, Appl
	41	12	60.0	90	12	US-10-023-523-34	Sequence 34, Appl
	42	12	60.0	93	10	US-09-998-598-2578	Sequence 2578, Ap
	43	12	60.0	99	10	US-09-969-373-880	Sequence 880, App
	44	11.8	59.0	45	10	US-09-742-693-2	Sequence 2, Appli
c	45	11.8	59.0	85	10	US-09-062-104A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-015-536-13/c
; Sequence 13, Application US/10015536
; Patent No. US20020165149A1
; GENERAL INFORMATION:
; APPLICANT: Kranz, David M.
; APPLICANT: Starwalt, Scott
; APPLICANT: Bluestone, Jeffrey A.
; TITLE OF INVENTION: Mutated Class II Major Histocompatibility Proteins
; FILE REFERENCE: 103-00
; CURRENT APPLICATION NUMBER: US/10/015,536
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: 60/254,248
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: nucleotide primer
US-10-015-536-13

Query Match 71.0%; Score 14.2; DB 9; Length 40;
Best Local Similarity 84.2%; Pred. No. 5.9e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 GACGAGGGCGAGCACGGCG 20
           |||| ||||||| || |||
Db      39 GACGTGGGCGAGTACCGCG 21
```

RESULT 2
US-09-908-153B-26
; Sequence 26, Application US/09908153B
; Patent No. US20020168714A1
; GENERAL INFORMATION:
; APPLICANT: Barbas, Carlos F.
; APPLICANT: Beeril, Roger
; APPLICANT: Schopfer, Ulrich

```
;  TITLE OF INVENTION: REGULATION OF GENE EXPRESSION USING
;  TITLE OF INVENTION:  SINGLE-CHAIN, MONOMERIC, LIGAND DEPENDENT POLYPEPTIDE
;  TITLE OF INVENTION:  SWITCHES
;  FILE REFERENCE: TSRI 725.1
;  CURRENT APPLICATION NUMBER: US/09/908,153B
;  CURRENT FILING DATE:  2001-07-18
;  PRIOR APPLICATION NUMBER: US 09/619,063
;  PRIOR FILING DATE: 2000-07-18
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
;   LENGTH: 45
;   TYPE: DNA
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Synthesized
US-09-908-153B-26

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 45;
  Best Local Similarity   84.2%;  Pred. No. 5.8e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGC 19
          ||| ||||||| || ||||
Db      3 GGAGGAGGGCGCGCCCGGC 21


RESULT 3
US-09-815-242-1169/c
; Sequence 1169, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1169
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-815-242-1169

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 60;
  Best Local Similarity   84.2%;  Pred. No. 5.7e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGC 19
          |||| ||||| |||||| |
Db     47 GGACCAGGGCCAGCACGCC 29


RESULT 4
US-09-815-242-1192/c
; Sequence 1192, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1192
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-815-242-1192

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 60;
  Best Local Similarity   84.2%;  Pred. No. 5.7e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGC 19
          |||| ||||| |||||| |
Db     47 GGACCAGGGCCAGCACGCC 29


RESULT 5
US-09-894-633A-29
; Sequence 29, Application US/09894633A
; Patent No. US20020124285A1
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy
;  APPLICANT:  Dubois, Patrice
;  APPLICANT:  Malven, Marianne
;  APPLICANT:  Masucci, James
;  TITLE OF INVENTION: PLANT REGULATORY SEQUENCES FOR SELECTIVE CONTROL OF GENE EXPRE
;  FILE REFERENCE: 38-21(15856)B
;  CURRENT APPLICATION NUMBER: US/09/894,633A
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: 60/214,357
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 09/894,633
;  PRIOR FILING DATE: 2000-06-28
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
;   LENGTH: 37
```

```
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(37)
;   OTHER INFORMATION: synthetic primer sequence
US-09-894-633A-29

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 37;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACGGCG 20
           ||||||||  |||  |||||
Db      14 GGACGAGGAGGAGGCCGGCG 33


RESULT 6
US-09-905-291A-89
; Sequence 89, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/905,291A
;  CURRENT FILING DATE:  2001-07-12
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 89
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: oligonucleotide probe
US-09-905-291A-89

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 50;
  Best Local Similarity   93.3%;  Pred. No. 1.3e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCAC 16
           ||||||||||| ||
Db      10 GACGAGGGCGAGTAC 24


RESULT 7
US-09-909-320-89
; Sequence 89, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/909,320
;  CURRENT FILING DATE:  2002-01-04
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
```

```
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 89
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: oligonucleotide probe
US-09-909-320-89

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 50;
  Best Local Similarity   93.3%;  Pred. No. 1.3e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCAC 16
           ||||||||||||| ||
Db      10 GACGAGGGCGAGTAC 24


RESULT 8
US-09-909-088B-89
; Sequence 89, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
;  APPLICANT: Genentech, Inc.
;  APPLICANT:  Ashkenazi, Avi
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan L.
;  APPLICANT:  Ferrara, Napoleone
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, A.
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, Christopher J.
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth, J.
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Mather, Jennie P.
;  APPLICANT:  Pan, James
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William, I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: 10466-14
;  CURRENT APPLICATION NUMBER: US/09/909,088B
;  CURRENT FILING DATE:  2001-07-18
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: US 60/143,048
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: US 60/145,698
;  PRIOR FILING DATE: 1999-07-26
;  PRIOR APPLICATION NUMBER: US 60/146,222
;  PRIOR FILING DATE: 1999-07-28
;  PRIOR APPLICATION NUMBER: PCT/US99/20594
;  PRIOR FILING DATE: 1999-09-08
;  PRIOR APPLICATION NUMBER: PCT/US99/20944
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: PCT/US99/21090
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/21547
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: PCT/US99/23089
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: PCT/US99/28214
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: PCT/US99/28313
;  PRIOR FILING DATE: 1999-11-30
;  PRIOR APPLICATION NUMBER: PCT/US99/28564
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28565
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US99/30095
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: PCT/US99/30911
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US99/30999
;  PRIOR FILING DATE: 1999-12-20
;  PRIOR APPLICATION NUMBER: PCT/US00/00219
;  PRIOR FILING DATE: 2000-01-05
;  NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 89
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: oligonucleotide probe
US-09-909-088B-89

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 50;
  Best Local Similarity   93.3%;  Pred. No. 1.3e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCAC 16
           ||||||||||||| ||
Db      10 GACGAGGGCGAGTAC 24


RESULT 9
US-09-772-656-4
; Sequence 4, Application US/09772656
; Patent No. US20010049832A1
; GENERAL INFORMATION:
;  APPLICANT:  Bruce, Wesley B.
;  TITLE OF INVENTION:  Root Transcriptional Factor
;  FILE REFERENCE:  1167R
;  CURRENT APPLICATION NUMBER: US/09/772,656
;  CURRENT FILING DATE:  2001-01-30
;  PRIOR APPLICATION NUMBER:  US 60/178,916
```

```
;  PRIOR FILING DATE:  2000-02-01
;  PRIOR APPLICATION NUMBER:  US Unknown
;  PRIOR FILING DATE:  2001-01-19
;  NUMBER OF SEQ ID NOS:  12
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Zea mays
US-09-772-656-4

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 27;
  Best Local Similarity   83.3%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGC 19
          | |||| ||||||| |||
Db      5 GGCGAGCGCGAGCAAGGC 22


RESULT 10
US-09-969-373-679/c
; Sequence 679, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
;  APPLICANT: Effertz, Roger J.
;  APPLICANT:  Hauge, Brian M.
;  TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
;  FILE REFERENCE: 38-10(52679)A
;  CURRENT APPLICATION NUMBER: US/09/969,373
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US 09/754,853
;  PRIOR FILING DATE: 2001-01-05
;  PRIOR APPLICATION NUMBER: US 09/760,427
;  PRIOR FILING DATE: 2001-01-13
;  PRIOR APPLICATION NUMBER: US 09/855,768
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 679
;   LENGTH: 96
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-969-373-679

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 96;
  Best Local Similarity   87.5%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACG 17
          ||||| | ||||||||
Db     71 GACGACGACGAGCACG 56


RESULT 11
US-09-780-053-15/c
; Sequence 15, Application US/09780053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
;  APPLICANT: Rene S. Hubert
;  APPLICANT:  Daniel E.H. Afar
;  APPLICANT:  Pia M. Challita-Eid
;  APPLICANT:  Mary Faris
;  APPLICANT:  Elana Levin
;  APPLICANT:  Steve Chappell Mitchell
;  APPLICANT:  Aya Jakobovits
;  TITLE OF INVENTION: 83P5G4:  A TISSUE SPECIFIC PROTEIN
;  TITLE OF INVENTION:  HIGHLY EXPRESSED IN PROSTATE CANCER
;  FILE REFERENCE: 129.5USU1
;  CURRENT APPLICATION NUMBER: US/09/780,053
;  CURRENT FILING DATE:  2001-02-09
;  PRIOR APPLICATION NUMBER: 60/181,261
;  PRIOR FILING DATE: 2000-02-09
;  NUMBER OF SEQ ID NOS: 716
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer
US-09-780-053-15

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          | ||| | ||||| |||||
Db     23 GTCGACGACGAGCGCGGCG 5


RESULT 12
US-09-779-308-15/c
; Sequence 15, Application US/09779308
; Patent No. US20020150972A1
; GENERAL INFORMATION:
;  APPLICANT: Mary Faris
;  APPLICANT:  Daniel E.H. Afar
;  APPLICANT:  Pia M. Challita-Eid
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Elana Levin
;  APPLICANT:  Steve Chappell Mitchell
;  APPLICANT:  Aya Jakobovits
;  TITLE OF INVENTION: 34P3D7:  A TISSUE SPECIFIC PROTEIN
;  TITLE OF INVENTION:  HIGHLY EXPRESSED IN PROSTATE CANCER
;  FILE REFERENCE: 129.4USU1
;  CURRENT APPLICATION NUMBER: US/09/779,308
;  CURRENT FILING DATE:  2001-02-08
;  PRIOR APPLICATION NUMBER: 60/181,020
;  PRIOR FILING DATE: 2000-02-08
;  NUMBER OF SEQ ID NOS: 718
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer
US-09-779-308-15

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 2.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          | ||| | ||||| |||||
Db     23 GTCGACGACGAGCGCGGCG 5


RESULT 13
US-09-852-922-6
; Sequence 6, Application US/09852922
; Patent No. US20020076768A1
; GENERAL INFORMATION:
;  APPLICANT: TOYO BOSEKI KABUSHIKI KAISHA
;  TITLE OF INVENTION: MODIFIED THERMOSTABLE DNA POLYMERASE
;  FILE REFERENCE: 000053
;  CURRENT APPLICATION NUMBER: US/09/852,922
;  CURRENT FILING DATE:  2001-05-10
;  PRIOR APPLICATION NUMBER: 2000-138796
;  PRIOR FILING DATE: 2000-05-11
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 34
;   TYPE: DNA
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-852-922-6

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 34;
  Best Local Similarity   78.9%;  Pred. No. 2.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           ||||||||||| |   ||
Db      13 GACGAGGGCGAGGAGTTCG 31


RESULT 14
US-09-852-922-7/c
; Sequence 7, Application US/09852922
; Patent No. US20020076768A1
; GENERAL INFORMATION:
;  APPLICANT: TOYO BOSEKI KABUSHIKI KAISHA
;  TITLE OF INVENTION: MODIFIED THERMOSTABLE DNA POLYMERASE
;  FILE REFERENCE: 000053
;  CURRENT APPLICATION NUMBER: US/09/852,922
;  CURRENT FILING DATE:  2001-05-10
;  PRIOR APPLICATION NUMBER: 2000-138796
;  PRIOR FILING DATE: 2000-05-11
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-852-922-7

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 34;
  Best Local Similarity   78.9%;  Pred. No. 2.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           ||||||||||| |   ||
Db      22 GACGAGGGCGAGGAGTTCG 4


RESULT 15
US-09-962-055-36
; Sequence 36, Application US/09962055
; Patent No. US20020052033A1
;   GENERAL INFORMATION:
;        APPLICANT: Lees, Ann M.
;                   Lees, Robert S.
;                   Law, Simon W.
;                   Arjona, Anibal A.
;        TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
;                            BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
;                            TREATING ATHEROSCLEROSIS
;        NUMBER OF SEQUENCES: 42
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fish & Richardson P.C.
;             STREET: 225 Franklin Street
;             CITY: Boston
;             STATE: MA
;             COUNTRY: USA
;             ZIP: 02110-2804
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/962,055
;             FILING DATE: 24-Sep-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/979,608
;             FILING DATE: 26-NOV-1997
;             APPLICATION NUMBER: US 60/031,930
;             FILING DATE: 27-NOV-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Myers, Louis
;             REGISTRATION NUMBER: 35,965
;             REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 617/542-5070
;             TELEFAX: 617/542-8906
;   INFORMATION FOR SEQ ID NO: 36:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 48 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        FEATURE:
;             NAME/KEY:  Coding Sequence
;             LOCATION:  1...48
;        SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-962-055-36

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 48;
  Best Local Similarity   78.9%;  Pred. No. 2.8e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GACGAGGGCGAGCACGGCG 20
           ||||||| |||  ||| ||
Db      28 GACGAGGACGACGACGACG 46


Search completed: November 23, 2002, 06:42:02
Job time : 18.25 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-3
Perfect score: 20
Sequence: 1 ggacgagggcgagcacggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	15.2	76.0	79	9	AI000040	AI000040 os46c07.s
c	2	14.8	74.0	87	10	AW064096	AW064096 SP0522 KR
c	3	14.8	74.0	91	17	AZ919805	AZ919805 1006016G0
c	4	14.8	74.0	91	17	BH218633	BH218633 1006080D1
c	5	14.4	72.0	64	9	AI900558	AI900558 sc07g10.y
c	6	14.4	72.0	84	17	BH226921	BH226921 1006135G0
c	7	14.2	71.0	80	10	AV920979	AV920979 AV920979
c	8	14.2	71.0	83	9	AI430740	AI430740 mh51e05.x
	9	14.2	71.0	89	10	BE264581	BE264581 601192351
	10	14.2	71.0	90	17	TA178C01P	AL474684 T. brucei
	11	13.8	69.0	79	17	BH407770	BH407770 1007050E0
c	12	13.8	69.0	80	12	BF022188	BF022188 uy47h01.y
	13	13.8	69.0	85	13	BI855354	BI855354 603382417
c	14	13.8	69.0	93	17	AZ848413	AZ848413 2M0149F24
c	15	13.8	69.0	96	17	BH416254	BH416254 1007047B0
	16	13.6	68.0	43	9	AA663711	AA663711 ae71a03.s
c	17	13.6	68.0	49	17	BH802890	BH802890 1008097B0
	18	13.6	68.0	50	9	AU103161	AU103161 AU103161
	19	13.6	68.0	50	9	AU103162	AU103162 AU103162
	20	13.6	68.0	50	9	AU103170	AU103170 AU103170
	21	13.6	68.0	50	9	AU103171	AU103171 AU103171
	22	13.6	68.0	50	9	AU103175	AU103175 AU103175
	23	13.6	68.0	50	9	AU103176	AU103176 AU103176
	24	13.6	68.0	50	9	AU103180	AU103180 AU103180
c	25	13.6	68.0	61	13	BI109390	BI109390 602897183
	26	13.6	68.0	67	10	AW246906	AW246906 2822525.5
c	27	13.6	68.0	74	13	BM097862	BM097862 EBem04_SQ
c	28	13.6	68.0	77	14	BQ760479	BQ760479 EBro02_SQ
c	29	13.6	68.0	79	14	BQ167594	BQ167594 WHE0077_H
	30	13.6	68.0	82	9	AI682177	AI682177 tp48e11.x
	31	13.6	68.0	82	13	BI755472	BI755472 603027324
	32	13.6	68.0	87	14	F37066	F37066 HSPD35102 H
c	33	13.6	68.0	89	17	BH414941	BH414941 1007040E0
c	34	13.6	68.0	95	10	AV913999	AV913999 AV913999
c	35	13.6	68.0	95	10	BE470627	BE470627 WHE0262_B
c	36	13.6	68.0	95	17	BH215460	BH215460 1006027B1
c	37	13.6	68.0	97	17	AZ771502	AZ771502 1M0573N08
	38	13.6	68.0	100	14	BM872771	BM872771 mgmy004xL
	39	13.6	68.0	100	14	BQ759154	BQ759154 EBma08_SQ
	40	13.4	67.0	66	17	BH641036	BH641036 1008039H0
c	41	13.2	66.0	69	17	AZ820652	AZ820652 2M0093B03
c	42	13.2	66.0	77	13	BI963211	BI963211 ie57g03.x
	43	13.2	66.0	79	14	R87701	R87701 yo44d11.r1
	44	13.2	66.0	85	17	BH621101	BH621101 1007104G0
	45	13.2	66.0	88	9	AA467645	AA467645 vf08c08.r

ALIGNMENTS

RESULT 1
AI000040
LOCUS AI000040 79 bp mRNA linear EST 05-JUN-1998
DEFINITION os46c07.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1608396 3' similar to TR:O15551 O15551 RAT VENTRAL PROSTATE.1 HOMOLOG. ;contains TAR1.b1 MSR1 repetitive element ;, mRNA sequence.
ACCESSION AI000040
VERSION AI000040.1 GI:3190594
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 79)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```
           Trace considered overall poor quality
           Seq primer: -40m13 fwd. ET from Amersham
           High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1608396"
                     /clone_lib="NCI_CGAP_Br2"
                     /sex="female, pooled"
                     /tissue_type="breast"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from pooled bulk
                     breast tumor tissue, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  This library is the normalized version of
                     NCI_CGAP_Br1.1. Library was constructed by Bento Soares
                     and M. Fatima Bonaldo. "
BASE COUNT       17 a     18 c     36 g      8 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 9;  Length 79;
  Best Local Similarity   85.0%;  Pred. No. 2.1e+04;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGCG 20
          ||||||||  ||||| ||||
Db     52 GGACGAGGATGAGCAGGGCG 71


RESULT 2
AW064096/c
LOCUS       AW064096                 87 bp    mRNA    linear   EST 07-DEC-2000
DEFINITION  SP0522 KRIBB Human CD4 intrathymic T-cell cDNA library Homo sapiens
            cDNA 3', mRNA sequence.
ACCESSION   AW064096
VERSION     AW064096.1  GI:8888033
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   Goh,S.-H., Park,J.-H., Lee,Y.J., Lee,H.G., Yoo,H.-S., Lee,I.-C.,
            Park,J.-H., Kim,Y.-S. and Lee,C.-C.
  TITLE     Gene expression profile and identification of differentially
            expressed transcripts during human intrathymic T-cell development
            by cDNA sequencing analysis
  JOURNAL   Genomics 70 (1), 1-18 (2000)
  MEDLINE   20541704
COMMENT     Contact: Sung-Ho Goh
            Genome Center
            Korea Research Institute of Bioscience and Biotechnology
            Oun-dong 52, Yu Sung-Gu, Daejon 305-333, Republic of Korea
            Tel: 82-42-860-4473
            Fax: 82-42-860-4479
            Email: gohsh@mail.kribb.re.kr
            Seq primer: T7
            High quality sequence stop: 87
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="KRIBB Human CD4 intrathymic T-cell cDNA
                     library"
                     /tissue_type="Thymus"
                     /cell_type="Intrathymic T-cell"
                     /dev_stage="CD3+4+8- single positive stage"
                     /note="Vector: pGEM-T; cDNA was made from total
                     cytoplasmic RNA of sorted human intrathymic CD3+4+8-
                     T-cell, adaptor ligated, amplified with PCR, and cloned
                     into pGEM-T vector."
BASE COUNT       17 a     38 c     14 g     18 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 10;  Length 87;
  Best Local Similarity   88.9%;  Pred. No. 3e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 ACGAGGGCGAGCACGGCG 20
          |||| ||||||| |||||
Db     38 ACGAAGGCGAGCGCGGCG 21


RESULT 3
AZ919805/c
LOCUS       AZ919805                 91 bp    DNA     linear   GSS 17-DEC-2001
DEFINITION  1006016G03.y1 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   AZ919805
VERSION     AZ919805.1  GI:13389882
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1006016  row: 41
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       25 a     27 c     22 g     17 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 17;  Length 91;
  Best Local Similarity   88.9%;  Pred. No. 3e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGG 18
          | ||||||||||| ||||
```

```
Db     34 GCACGAGGGCGAGTACGG 17


RESULT 4
BH218633/c
LOCUS       BH218633                  91 bp    DNA     linear   GSS 08-NOV-2001
DEFINITION  1006080D10.xl 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   BH218633
VERSION     BH218633.1  GI:16811673
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site found so sequence was trimmed.
            Post-ligation sequence submitted separately.
            Plate: 1006080  row: 28
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT        7 a     22 c     33 g     29 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 17;  Length 91;
  Best Local Similarity   88.9%;  Pred. No. 3e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGC 19
          ||||||| |||||||| |
Db     77 GACGAGGTCGAGCACGAC 60


RESULT 5
AI900558/c
LOCUS       AI900558                  64 bp    mRNA    linear   EST 30-NOV-2001
DEFINITION  sc07g10.y1 Gm-c1012 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1012-1531 5' similar to SW:PSAF_FLATR P46486 PHOTOSYSTEM I
            REACTION CENTRE SUBUNIT III PRECURSOR ;, mRNA sequence.
ACCESSION   AI900558
VERSION     AI900558.1  GI:5606404
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand This clone is
            available through: ResGen, Invitrogen Corp. 2130 South Memorial
            Parkway Huntsville, AL 35801 For further information call: (800
            )-533-4363 or contact via email: ccu@resgen.com
            Seq primer: -40RP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1012-1531"
                     /clone_lib="Gm-c1012"
                     /tissue_type="Apical shoot tips, 9-10 day old etiolated
                     seedlings"
                     /lab_host="XL10-Gold"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from the apical shoots of 9 to 10 day old etiolated
                     seedlings. The shoot tips including any emerged leaves
                     were harvested for mRNA isolation. The cDNA library was
                     prepared using the Stratagene pBluescript II XR cDNA
                     library construction kit. Complementary DNA was
                     synthesized from mRNA using a primer consisting of a poly
                     (dT) sequence with a XhoI restriction site. EcoRI adapters
                     were ligated to the blunt-ended cDNA fragments followed by
                     XhoI digestion. The cDNA fragments were directionally
                     cloned into the EcoRI-XhoI restriction site of the
                     pBluescript vector. The ligated cDNA fragments were
                     transformed into XL10-Gold host cells. This library was
                     constructed by Dr. Randy Shoemaker and Dr. John
                     Erpelding."
BASE COUNT        4 a     34 c     11 g     15 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 9;  Length 64;
  Best Local Similarity   93.8%;  Pred. No. 4.2e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 GAGGGCGAGCACGGCG 20
          |||||| |||||||||
Db     34 GAGGGCCAGCACGGCG 19


RESULT 6
BH226921/c
LOCUS       BH226921                  84 bp    DNA     linear   GSS 08-NOV-2001
DEFINITION  1006135G08.2EL_y1 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   BH226921
VERSION     BH226921.1  GI:16826390
```

```
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 84)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1006135  column: 32
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       11 a     33 c     20 g     20 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 17;  Length 84;
  Best Local Similarity   93.8%;  Pred. No. 4.2e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACG 17
          |||||||||||| |||
Db     69 GACGAGGGCGAGGACG 54


RESULT 7
AV920979/c
LOCUS       AV920979                 80 bp    mRNA    linear   EST 18-JAN-2002
DEFINITION  AV920979 K. Sato unpublished cDNA library, cv. Haruna Nijo
            germination shoots Hordeum vulgare subsp. vulgare cDNA clone
            bags16f09 3', mRNA sequence.
ACCESSION   AV920979
VERSION     AV920979.1  GI:18216758
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare.
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 80)
  AUTHORS   Sato,K., Saisho,D. and Takeda,K.
  TITLE     Barley EST sequencing project in NIG and Okayama Univ
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="Hordeum vulgare subsp. vulgare"
                     /cultivar="Haruna Nijo"
                     /db_xref="taxon:112509"
                     /clone="bags16f09"
                     /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
                     Nijo germination shoots"
                     /tissue_type="shoots"
                     /dev_stage="germination"
BASE COUNT       12 a     31 c     15 g     22 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 80;
  Best Local Similarity   84.2%;  Pred. No. 5.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          |||||||||||  | ||||
Db     48 GACGAGGGCGACGAGGGCG 30


RESULT 8
AI430740/c
LOCUS       AI430740                 83 bp    mRNA    linear   EST 09-MAR-1999
DEFINITION  mh51e05.x1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone IMAGE:446048 3', mRNA sequence.
ACCESSION   AI430740
VERSION     AI430740.1  GI:4276576
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 83)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end
            Possible reversed clone: polyT not found.
FEATURES             Location/Qualifiers
     source          1. .83
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:446048"
                     /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                     /sex="unknown"
                     /tissue_type="placenta"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
```

```
                [5'
                TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                (Pharmacia), digested with Not I and cloned into the Not I
                and Eco RI sites of the modified pT7T3 vector. Library
                went through one round of normalization, and was
                constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT        9 a     38 c     23 g     13 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 83;
  Best Local Similarity   84.2%;  Pred. No. 5.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGC 19
          ||| ||||  |||||||||
Db     26 GGAGGAGGCGGAGCACGGC 8


RESULT 9
BE264581
LOCUS       BE264581                 89 bp    mRNA    linear   EST 13-JUL-2000
DEFINITION  601192351F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3536245 5',
            mRNA sequence.
ACCESSION   BE264581
VERSION     BE264581.1  GI:9138138
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 89)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Plate: LLCM218  row: e  column: 14
            High quality sequence stop: 89.
FEATURES             Location/Qualifiers
     source          1. .89
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3536245"
                     /clone_lib="NIH_MGC_7"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT       19 a     26 c     28 g     16 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 89;
  Best Local Similarity   84.2%;  Pred. No. 5.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACGGC 19
          ||||||| |||||  ||||
Db     52 GGACGAGTGCGAGTCCGGC 70


RESULT 10
TA178C01P
LOCUS       TA178C01P                90 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 178c01, forward sequence,
            genomic survey sequence.
ACCESSION   AL474684
VERSION     AL474684.1  GI:11839870
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="178c01"
BASE COUNT       36 a     22 c     23 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 90;
  Best Local Similarity   84.2%;  Pred. No. 5.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GACGAGGGCGAGCACGGCG 20
          ||||| ||| |||| ||||
Db     38 GACGAAGGCAAGCATGGCG 56


RESULT 11
BH407770
LOCUS       BH407770                 79 bp    DNA     linear   GSS 12-DEC-2001
DEFINITION  1007050E04.1EL_y1 1007 - RescueMu Grid H Zea mays genomic, DNA
            sequence.
ACCESSION   BH407770
VERSION     BH407770.1  GI:17572739
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site of ends cut by single endonuclease.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1007050  column: 25
```

```
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1007 - RescueMu Grid H"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units.  For more information on RescueMu, go to the  web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.'  Grid H was grown at Berkeley in 2001.  DNA
                     was extracted from leaf punches,  double digested using
                     BamHI and BglII, and ligated to form circular plasmids.
                     DH10B cells were transformed and then screened on LB
                     plates with ampicillin."
BASE COUNT       13 a     22 c     36 g      8 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 79;
  Best Local Similarity   88.2%;  Pred. No. 7.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CGAGGGCGAGCACGGCG 20
          |||||||||| ||| ||
Db     62 CGAGGGCGAGGACGTCG 78


RESULT 12
BF022188/c
LOCUS       BF022188                 80 bp    mRNA    linear   EST 29-DEC-2000
DEFINITION  uy47h01.y1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:3662737 5',
            mRNA sequence.
ACCESSION   BF022188
VERSION     BF022188.1  GI:10753520
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 80)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml

            MGI:1423505
            Seq primer: -40RP from Gibco
            High quality sequence stop: 70.
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="Mus musculus"
                     /strain="CZECH II"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3662737"
                     /clone_lib="NCI_CGAP_Lu30"
                     /tissue_type="tumor, metastatic to mammary"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; transgenic model WNT-1, expression driven by
                     MMTV-LTR enhancer; Cloned unidirectionally.  Primer: Oligo
                     dT.   Library constructed by Life Technologies.
                     Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT        6 a     31 c     19 g     24 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 12;  Length 80;
  Best Local Similarity   88.2%;  Pred. No. 7.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GGACGAGGGCGAGCACG 17
          |||||||| |||| |||
Db     66 GGACGAGGACGAGGACG 50


RESULT 13
BI855354
LOCUS       BI855354                 85 bp    mRNA    linear   EST 10-OCT-2001
DEFINITION  603382417F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5414669 5',
            mRNA sequence.
ACCESSION   BI855354
VERSION     BI855354.1  GI:15996101
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12058  row: a  column: 06
            High quality sequence stop: 85.
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5414669"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT        7 a     38 c     31 g      9 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 13;  Length 85;
  Best Local Similarity   88.2%;  Pred. No. 7.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CGAGGGCGAGCACGGCG 20
          || ||||| ||||||||
Db      7 CGTGGGCGCGCACGGCG 23


RESULT 14
AZ848413/c
```

```
LOCUS       AZ848413                93 bp    DNA     linear   GSS 21-FEB-2001
DEFINITION  2M0149F24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0149F24 F, DNA sequence.
ACCESSION   AZ848413
VERSION     AZ848413.1  GI:13031481
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0149  row: F  column: 24
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 93.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0149F24"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       14 a     35 c     28 g     16 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 93;
  Best Local Similarity   88.2%;  Pred. No. 7.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 ACGAGGGCGAGCACGGC 19
           ||| ||||| |||||||
Db      28 ACGGGGGCGCGCACGGC 12


RESULT 15
BH416254/c
```

```
LOCUS       BH416254                96 bp    DNA     linear   GSS 12-DEC-2001
DEFINITION  1007047B09.x1 1007 - RescueMu Grid H Zea mays genomic, DNA
            sequence.
ACCESSION   BH416254
VERSION     BH416254.1  GI:17598453
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1007047  column: 14
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1007 - RescueMu Grid H"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units.  For more information on RescueMu, go to the  web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.'  Grid H was grown at Berkeley in 2001.  DNA
                     was extracted from leaf punches,  double digested using
                     BamHI and BglII, and ligated to form circular plasmids.
                     DH10B cells were transformed and then screened on LB
                     plates with ampicillin."
BASE COUNT       14 a     41 c     25 g     16 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 96;
  Best Local Similarity   88.2%;  Pred. No. 7.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GGACGAGGGCGAGCACG 17
           ||||||||||||| | |
Db      17 GGACGAGGGCGAGAATG 1


Search completed: November 26, 2002, 04:06:38
Job time : 766.8 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-4
Perfect score: 20
Sequence: 1 tgggtccggagcctgaatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		15.2	76.0	82	6	I04095	I04095 Sequence 10
c 2		14.2	71.0	42	6	AR194079	AR194079 Sequence
3		14.2	71.0	66	6	AX112901	AX112901 Sequence
c 4		14.2	71.0	77	12	AF401224	AF401224 Synthetic
5		13.8	69.0	20	6	I86696	I86696 Sequence 22
6		13.6	68.0	24	6	E13049	E13049 Primer. 4/1
c 7		13.6	68.0	96	6	AX049355	AX049355 Sequence
8		13.4	67.0	21	6	AX034723	AX034723 Sequence
c 9		13.2	66.0	23	6	AX251733	AX251733 Sequence
10		13.2	66.0	74	6	AX478566	AX478566 Sequence
c 11		13.2	66.0	82	6	AX478567	AX478567 Sequence
c 12		13.2	66.0	84	8	PVTRNL2	X08010 Phaseolus v
13		12.8	64.0	17	6	AX017782	AX017782 Sequence
c 14		12.8	64.0	19	6	AX132324	AX132324 Sequence
15		12.8	64.0	38	6	I21500	I21500 Sequence 47
16		12.8	64.0	66	6	A07506	A07506 Nucleotide
17		12.8	64.0	91	9	HUMCP2106	M23225 Human mutan
c 18		12.6	63.0	19	6	AR148676	AR148676 Sequence
19		12.6	63.0	23	6	AX306373	AX306373 Sequence
20		12.6	63.0	23	6	AX320086	AX320086 Sequence
21		12.6	63.0	23	6	AX352397	AX352397 Sequence
22		12.6	63.0	23	6	AX419617	AX419617 Sequence
23		12.6	63.0	24	6	E26146	E26146 Novel brain
24		12.6	63.0	25	6	AX127099	AX127099 Sequence
25		12.6	63.0	32	6	E34282	E34282 Environment
26		12.6	63.0	32	6	E34338	E34338 Environment
c 27		12.6	63.0	33	6	AX326729	AX326729 Sequence
28		12.6	63.0	33	6	AX326730	AX326730 Sequence
c 29		12.6	63.0	40	6	AX456208	AX456208 Sequence
30		12.6	63.0	40	6	AX456258	AX456258 Sequence
c 31		12.6	63.0	41	6	AX268538	AX268538 Sequence
32		12.6	63.0	49	6	AX279731	AX279731 Sequence
c 33		12.6	63.0	51	6	AX161788	AX161788 Sequence
34		12.6	63.0	63	6	AX468790	AX468790 Sequence
35		12.6	63.0	63	6	AX468800	AX468800 Sequence
c 36		12.6	63.0	69	6	AR081748	AR081748 Sequence
c 37		12.6	63.0	69	6	AR167980	AR167980 Sequence
38		12.6	63.0	76	1	MCTRK2	X16757 Mycoplasma
39		12.6	63.0	78	10	U96708	U96708 Mus musculu
40		12.6	63.0	94	6	AX088803	AX088803 Sequence
41		12.6	63.0	97	6	AX106283	AX106283 Sequence
42		12.6	63.0	97	6	AX140574	AX140574 Sequence
43		12.6	63.0	97	6	AX200434	AX200434 Sequence
44		12.6	63.0	97	6	AX267090	AX267090 Sequence
45		12.4	62.0	21	6	AX034726	AX034726 Sequence

ALIGNMENTS

RESULT 1
I04095
LOCUS I04095 82 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 10 from Patent EP 0128733.
ACCESSION I04095
VERSION I04095.1 GI:591892
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 82)
AUTHORS Lee,J.M. and Ullrich,A.
TITLE Human insulin-like growth factor (IGF) and human epidermal growth factor (EGF) produced from a recombinant host, process, expression vector and recombinant host therefor, and IGF-containing

```
          pharmaceutical composition
  JOURNAL   Patent: EP 0128733-A1 10 19-DEC-1984;
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="unknown"
BASE COUNT       20 a     19 c     19 g     24 t
ORIGIN

  Query Match           76.0%;  Score 15.2;  DB 6;  Length 82;
  Best Local Similarity   85.0%;  Pred. No. 6.9e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAATCA 20
          ||||| |||| ||| |||||
Db     58 TGGGTACGGATCCTCAATCA 77


RESULT 2
AR194079/c
LOCUS       AR194079                  42 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 8 from patent US 6348333.
ACCESSION   AR194079
VERSION     AR194079.1  GI:20240671
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Niwa,M., Okamoto,M., Matsumoto,T. and Segawa,T.
  TITLE     VEGF-binding KDR polypeptide
  JOURNAL   Patent: US 6348333-A 8 19-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="unknown"
BASE COUNT       12 a     11 c      8 g     11 t
ORIGIN

  Query Match           71.0%;  Score 14.2;  DB 6;  Length 42;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAATC 19
          || ||||||||||  ||||
Db     29 TGAGTCCGGAGCCCAAATC 11


RESULT 3
AX112901
LOCUS       AX112901                  66 bp    DNA     linear   PAT 01-MAY-2001
DEFINITION  Sequence 3 from Patent WO0127294.
ACCESSION   AX112901
VERSION     AX112901.1  GI:13939380
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Devico,A.L., Fouts,T.R. and Tuskan,R.G.
  TITLE     Virus coat protein/receptor chimeras and methods of use
  JOURNAL   Patent: WO 0127294-A 3 19-APR-2001;
            UNIVERSITY OF MARYLAND BIOTECHNOLOGY INSTITUTE (US)
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Artificial Sequence - primer"
BASE COUNT       10 a     28 c     20 g      8 t
ORIGIN

  Query Match           71.0%;  Score 14.2;  DB 6;  Length 66;
  Best Local Similarity   84.2%;  Pred. No. 2.2e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGGTCCGGAGCCTGAATCA 20
          |||||||||||| ||  ||
Db      1 GGGTCCGGAGCCCGAGCCA 19


RESULT 4
AF401224/c
LOCUS       AF401224                  77 bp    DNA     linear   SYN 26-AUG-2001
DEFINITION  Synthetic construct GCN4F cis element sequence.
ACCESSION   AF401224
VERSION     AF401224.1  GI:15290628
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Xian,L., Quanhong,Y., Rihe,P., Aisheng,X. and Huiqin,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-2001) Shanghai Yong Ye Agro-Bioengineering Co.,
            Ltd, Shanghai Academy of Agricultural Sciences, Beidi Road 2901,
            Shanghai 201106, China
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="cis element containing five copies of the GCN4F
                     element; can be inserted into lacZ reporter plasmid for
                     yeast one-hybrid system"
BASE COUNT       21 a     19 c     21 g     16 t
ORIGIN

  Query Match           71.0%;  Score 14.2;  DB 12;  Length 77;
  Best Local Similarity   84.2%;  Pred. No. 2.2e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGGTCCGGAGCCTGAATCA 20
          ||| ||||| ||||| |||
Db     77 GGGCCCGGATCCTGACTCA 59


RESULT 5
I86696
LOCUS       I86696                    20 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 22 from patent US 5702895.
ACCESSION   I86696
VERSION     I86696.1  GI:3206414
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Matsunaga,H., Tsukumo,K., Wakisaka,S. and Yamane,A.
  TITLE     Method and kit for detecting methicillin-resistant Staphylococcus
            aureus
  JOURNAL   Patent: US 5702895-A 22 30-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        5 a      6 c      4 g      5 t
ORIGIN

  Query Match           69.0%;  Score 13.8;  DB 6;  Length 20;
  Best Local Similarity   88.2%;  Pred. No. 3.8e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GTCCGGAGCCTGAATCA 20
          ||||| | |||||||||
Db      2 GTCCGTAACCTGAATCA 18


RESULT 6
E13049
LOCUS       E13049                    24 bp    DNA     linear   PAT 27-APR-1998
```

```
DEFINITION  Primer.
ACCESSION   E13049
VERSION     E13049.1  GI:3251861
KEYWORDS    JP 1997121869-A/6.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Kondo,M. and Sudou,S.
  TITLE     MEA-2 GENE
  JOURNAL   Patent: JP 1997121869-A 6 13-MAY-1997;
            ITO HAM KK
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1997121869-A/6
            PD   13-MAY-1997
            PF   07-NOV-1995 JP 1995311638
            PI   KONDO MASAAKI, SUDOU SHIZUYO
            PC   C12N15/09,C07H21/02,C07H21/04,C07K14/47,C07K16/18,C12N1/21, PC
              C12Q1/68,
            PC   G01N33/53,G01N33/53,G01N33/531,(C12N1/21,C12R1:19); CC
            strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .24
            FT                   /organism='Artificial sequences'.
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        5 a      5 c      9 g      5 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 24;
  Best Local Similarity   80.0%;  Pred. No. 4.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAATCA 20
          | || ||||| || ||||||
Db      1 TTGGGCCGGAACCGGAATCA 20


RESULT 7
AX049355/c
LOCUS       AX049355                96 bp    DNA     linear   PAT 12-JAN-2001
DEFINITION  Sequence 25 from Patent WO0071709.
ACCESSION   AX049355
VERSION     AX049355.1  GI:12226111
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Giros,B., Gasnier,B., Sagne,C., el Mestikawy,S. and Hamon,M.
  TITLE     Polypeptides, vesicular carriers of glutamate and gaba
  JOURNAL   Patent: WO 0071709-A 25 30-NOV-2000;
            INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE; (INSERM)
            (FR)
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       23 a     23 c     26 g     24 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 96;
  Best Local Similarity   80.0%;  Pred. No. 4.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAATCA 20
          ||| ||||||||| | | ||
Db     81 TGGTTCCGGAGCCAGGAGCA 62


RESULT 8
AX034723
LOCUS       AX034723                21 bp    DNA     linear   PAT 15-NOV-2000
DEFINITION  Sequence 5 from Patent WO0052203.
ACCESSION   AX034723
VERSION     AX034723.1  GI:11190682
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Anthony,R.M., Brown,T.J. and French,G.L.
  TITLE     Identification of bacteria
  JOURNAL   Patent: WO 0052203-A 5 08-SEP-2000;
            ANTHONY RICHARD MICHAEL (GB) ; BROWN TIMOTHY JAMES (GB) ; KING S
            COLLEGE LONDON (GB) ; FRENCH GARY LAWRENCE (GB) ; GUY S & ST THOMAS
            S NATIONAL H (GB)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="OLIGO"
BASE COUNT        5 a      5 c      6 g      5 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 21;
  Best Local Similarity   93.3%;  Pred. No. 6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 CCGGAGCCTGAATCA 20
          || ||||||||||||
Db      1 CCAGAGCCTGAATCA 15


RESULT 9
AX251733/c
LOCUS       AX251733                23 bp    DNA     linear   PAT 05-OCT-2001
DEFINITION  Sequence 11 from Patent WO0168859.
ACCESSION   AX251733
VERSION     AX251733.1  GI:15985090
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Jing,S.
  TITLE     Il-17 receptor like molecules and uses thereof
  JOURNAL   Patent: WO 0168859-A 11 20-SEP-2001;
            Amgen Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer 2432-36"
BASE COUNT        4 a      10 c      5 g      4 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 23;
  Best Local Similarity   83.3%;  Pred. No. 7.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGGTCCGGAGCCTGAATC 19
          | |||||||| ||| |||
Db     20 GAGTCCGGAGGCTGGATC 3


RESULT 10
AX478566
LOCUS       AX478566                74 bp    DNA     linear   PAT 12-AUG-2002
```

```
DEFINITION  Sequence 16 from Patent WO0216645.
ACCESSION   AX478566
VERSION     AX478566.1  GI:22217335
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Brenner,S.
  TITLE     Polymorphic dna fragments and uses thereof
  JOURNAL   Patent: WO 0216645-A 16 28-FEB-2002;
            LYNX THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthesized oligonucleotide"
BASE COUNT       19 a     18 c     20 g     17 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 74;
  Best Local Similarity   83.3%;  Pred. No. 7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATC 19
           ||| ||||| ||||| ||
Db      48 GGGCCCGGATCCTGACTC 65


RESULT 11
AX478567/c
LOCUS       AX478567                  82 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 17 from Patent WO0216645.
ACCESSION   AX478567
VERSION     AX478567.1  GI:22217336
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Brenner,S.
  TITLE     Polymorphic dna fragments and uses thereof
  JOURNAL   Patent: WO 0216645-A 17 28-FEB-2002;
            LYNX THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthesized oligonucleotide"
BASE COUNT       19 a     22 c     20 g     21 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 82;
  Best Local Similarity   83.3%;  Pred. No. 7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATC 19
           ||| ||||| ||||| ||
Db      31 GGGCCCGGATCCTGACTC 14


RESULT 12
PVTRNL2/c
LOCUS       PVTRNL2                   84 bp    tRNA    linear   PLN 02-AUG-1994
DEFINITION  Phaseolus vulgaris mitochondrial tRNA-Leu(2).
ACCESSION   X08010
VERSION     X08010.1  GI:21056
KEYWORDS    mitochondrial import; transfer RNA-Leu.
SOURCE      Phaseolus vulgaris.
  ORGANISM  Phaseolus vulgaris
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Phaseolus.
REFERENCE   1  (bases 1 to 84)
  AUTHORS   Marechal-Drouard,L., Weil,J.H. and Guillemaut,P.
  TITLE     Import of several tRNAs from the cytoplasm into the mitochondria in
            bean Phaseolus vulgaris
  JOURNAL   Nucleic Acids Res. 16 (11), 4777-4788 (1988)
  MEDLINE   88262475
   PUBMED   3387208
COMMENT     sequence corresponds to mt tRNA-Leu2 and its cytoplasmic
            counterpart
            Data kindly reviewed (23-AUG-1988) by GUILLEMAUT P.
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="Phaseolus vulgaris"
                     /db_xref="taxon:3885"
     tRNA            1. .84
                     /product="tRNA-Leu"
                     /note="tRNA-Leu(2)"
     modified_base   10
                     /note="2-methylguanosine"
                     /mod_base=m2g
     modified_base   12
                     /note="4-acetylcytidine"
                     /mod_base=ac4c
     modified_base   16
                     /note="dihydrouridine"
                     /mod_base=d
     misc_feature    17
                     /note="unmodified in cyt tRNA-Leu"
     modified_base   17
                     /note="2'-O-methylguanosine"
                     /mod_base=gm
     modified_base   19
                     /note="dihydrouridine"
                     /mod_base=d
     modified_base   21
                     /note="acp3U"
                     /mod_base=OTHER
     modified_base   27
                     /note="2,2-dimethylguanosine"
                     /mod_base=m22g
     modified_base   33
                     /note="pseudouridine"
                     /mod_base=p
     modified_base   35
                     /note="unidentfied modified nucleotide"
                     /mod_base=OTHER
     modified_base   38
                     /note="1-methylguanosine"
                     /mod_base=m1g
     modified_base   40
                     /note="pseudouridine"
                     /mod_base=p
     modified_base   45
                     /note="2'-O-methyluridine"
                     /mod_base=um
     modified_base   52
                     /note="unidentfied modified nucleotide"
                     /mod_base=OTHER
     modified_base   56
                     /note="5-methylcytidine"
                     /mod_base=m5c
     modified_base   62
                     /note="5-methyluridine"
                     /mod_base=t
     modified_base   63
                     /note="pseudouridine"
                     /mod_base=p
     modified_base   66
                     /note="1-methyladenosine"
                     /mod_base=m1a
BASE COUNT       18 a     20 c     26 g     18 t      2 others
ORIGIN
```

```
  Query Match             66.0%;  Score 13.2;  DB 8;  Length 84;
  Best Local Similarity   78.9%;  Pred. No. 7e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAATC 19
          | || || |||||| ||||
Db     49 TCGGACCAGAGCCTNAATC 31


RESULT 13
AX017782
LOCUS       AX017782                 17 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 11 from Patent WO9946404.
ACCESSION   AX017782
VERSION     AX017782.1  GI:10042389
KEYWORDS    .
SOURCE      Hordeum vulgare.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 17)
  AUTHORS   Ramsey,L.D., Powell,W., Waugh,R., Swanston,J.S. and Thomas,W.T.
  TITLE     Dna sequences and their use for the selection of cereals
  JOURNAL   Patent: WO 9946404-A 11 16-SEP-1999;
            RAMSEY LUKE DOUGLAS (GB); SCOTTISH CROP RESEARCH INST (GB); POWELL
            WAYNE (GB); WAUGH ROBERT (GB); SWANSTON JOHN STUART (GB); THOMAS
            WILLIAM THEODORE BLAYNE (GB)
FEATURES             Location/Qualifiers
     source          1. .17
                     /organism="Hordeum vulgare"
                     /db_xref="taxon:4513"
BASE COUNT        4 a      4 c      6 g      3 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 17;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 GGGTCCGGAGCCTGAA 17
          ||||| | ||||||||
Db      1 GGGTCTGAAGCCTGAA 16


RESULT 14
AX132324/c
LOCUS       AX132324                 19 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 3542 from Patent WO0130362.
ACCESSION   AX132324
VERSION     AX132324.1  GI:14138629
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Robbins,J.M. and Tritz,R.
  TITLE     Ribozyme therapy for the treatment of proliferative skin and eye
            diseases
  JOURNAL   Patent: WO 0130362-A 3542 03-MAY-2001;
            IMMUSOL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Cdc25 hs ribozyme binding site"
BASE COUNT        3 a      7 c      6 g      3 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 19;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGA 16
          ||||||| ||||| ||
Db     18 TGGGTCCTGAGCCAGA 3


RESULT 15
I21500
LOCUS       I21500                   38 bp    DNA     linear   PAT 07-OCT-1996
DEFINITION  Sequence 47 from patent US 5521300.
ACCESSION   I21500
VERSION     I21500.1  GI:1601854
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Shah,J.S., Nietupski,R.M. and Liu,J.
  TITLE     Oligonucleotides complementary to mycobacterial nucleic acids
  JOURNAL   Patent: US 5521300-A 47 28-MAY-1996;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
BASE COUNT       10 a     11 c     15 g      2 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 38;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GTCCGGAGCCTGAATC 19
          | || ||||||||||
Db     23 GCCCAGAGCCTGAATC 38


Search completed: December  3, 2002, 18:12:18
Job time : 311.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-4
Perfect score: 20
Sequence: 1 tgggtccggagcctgaatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

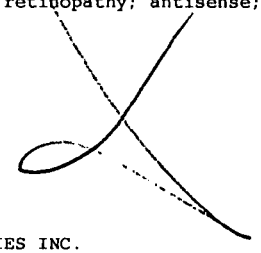
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	21	AAZ31434	Human neuropilin m
c 2		14.8	74.0	30	18	AAT68645	A 3' primer for th
c 3		14.2	71.0	42	19	AAV50375	PCR primer for KDR
4		14.2	71.0	66	22	AAH19643	Human immunodefici
c 5		13.8	69.0	51	22	AAL27594	Human SNP oligonuc
c 6		13.8	69.0	60	24	ABN35552	Human spliced tran
7		13.6	68.0	24	18	AAT74039	Male-enhanced anti
8		13.6	68.0	28	20	AAX01737	Rat GDNFR-beta cDN
9		13.6	68.0	65	24	ABN55425	Mouse spliced tran
c 10		13.6	68.0	65	24	ABN56524	Mouse spliced tran
c 11		13.6	68.0	96	23	ABL57923	Human VG51 exon 4.
12		13.4	67.0	21	21	AAA94347	Oligonucleotide 1c
c 13		13.2	66.0	23	22	AAF28284	Oligonucleotide SP
c 14		13.2	66.0	23	23	AAS15354	pCMV/SPORT vector
15		13.2	66.0	47	21	AAZ69457	Human map-related
16		13.2	66.0	50	22	AAL30919	Human SNP oligonuc
17		13.2	66.0	50	22	AAL30920	Human SNP oligonuc
18		13.2	66.0	50	22	AAL34469	Human SNP oligonuc
c 19		13.2	66.0	60	24	ABN36396	Human spliced tran
20		13.2	66.0	74	21	AAA64280	Oligonucleotide S-
21		13.2	66.0	74	21	AAA57120	Sequence tag plasm
22		13.2	66.0	74	21	AAA14490	Oligonucleotide S-
23		13.2	66.0	74	24	ABL53615	Oligonucleotide S-
c 24		13.2	66.0	82	21	AAA64281	Oligonucleotide S-
c 25		13.2	66.0	82	21	AAA57121	Sequence tag plasm
c 26		13.2	66.0	82	21	AAA14491	Oligonucleotide S-
c 27		13.2	66.0	82	22	ABA69723	Human foetal liver
c 28		13.2	66.0	82	22	ABA35254	Probe #13720 for g
c 29		13.2	66.0	82	22	AAK16622	Human brain expres
c 30		13.2	66.0	82	22	AAK43784	Human bone marrow
c 31		13.2	66.0	82	22	AAI24561	Probe #14494 for g
c 32		13.2	66.0	82	22	AAI48460	Probe #17146 used
c 33		13.2	66.0	82	24	ABS16436	Human genome-deriv
c 34		13.2	66.0	82	24	ABL53616	Oligonucleotide S-
35		12.8	64.0	17	20	AAZ20451	PCR primer Bmac399
36		12.8	64.0	18	21	AAZ44752	Human FADD primer
c 37		12.8	64.0	19	21	AAA85956	Cdc 25 hs ribozyme
c 38		12.8	64.0	19	22	AAH61118	Cdc25 hs ribozyme
39		12.8	64.0	23	24	ABK37920	Olfactory receptor
40		12.8	64.0	24	24	ABK52448	Human very low den
41		12.8	64.0	25	21	AAA51171	Primer for olfacto
42		12.8	64.0	25	21	AAA51172	Primer for olfacto
43		12.8	64.0	25	22	AAS13319	Human olfactory re
44		12.8	64.0	25	22	AAS13321	Human olfactory re
45		12.8	64.0	25	24	ABK37918	Olfactory receptor

ALIGNMENTS

RESULT 1
AAZ31434
ID AAZ31434 standard; DNA; 20 BP.
XX
AC AAZ31434;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3604.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -



```
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 4 A; 5 C; 7 G; 4 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAATCA 20
           ||||||||||||||||||||
Db       1 TGGGTCCGGAGCCTGAATCA 20


RESULT 2
AAT68645/c
ID   AAT68645 standard; DNA; 30 BP.
XX
AC   AAT68645;
XX
DT   14-AUG-1997  (first entry)
XX
DE   A 3' primer for the reverse tetracycline repressor gene.
XX
KW   DNA integration; LexA; gene therapy; oncogene inactivation;
KW   retrovirus; HIV-1; integrase; reverse tetracycline repressor;
KW   rtet gene; primer; PCR; polymerase chain reaction; ss.
XX
OS   Synthetic.
XX
PN   WO9720038-A1.
XX
PD   05-JUN-1997.
XX
PF   27-NOV-1996;   96WO-US19277.
XX
PR   01-DEC-1995;   95US-0008263.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Chow SA,  Goulaouic H;
XX
DR   WPI; 1997-310594/28.
XX
PT   Fusion protein comprising retroviral integrase and DNA binding
PT   regions - useful for integrating donor DNA into specific sites on
PT   target DNA, especially for inactivating oncogene(s) or inserting
PT   therapeutic genes
XX
PS   Example 10; Page 56; 92pp; English.
XX
CC   A 3' primer (AAT68645) containing a BamHI site and a 5' primer
CC   (AAT68644) containing a KpnI site were used in the PCR amplification
CC   of the E. coli reverse tetracycline repressor (rTET) gene.  The
CC   PCR product was used to prepare a construct coding for an HIV-1
CC   integrase-rTET fusion.  The activities of this fusion protein
CC   confirmed the generality of a method of fusing integrases from
CC   different sources, e.g. HIV-1 and FIV, with different DNA
CC   binding proteins, e.g. LexA and rTet, to achieve site-directed
CC   integration into a target DNA molecule (see also AAW17781-84).
XX
SQ   Sequence 30 BP; 7 A; 10 C; 5 G; 8 T; 0 other;

  Query Match             74.0%;  Score 14.8;  DB 18;  Length 30;
  Best Local Similarity   88.9%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAAT 18
           |||||||||| ||||| |
Db      19 TGGGTCCGGATCCTGACT 2


RESULT 3
AAV50375/c
ID   AAV50375 standard; DNA; 42 BP.
XX
AC   AAV50375;
XX
DT   11-NOV-1998  (first entry)
XX
DE   PCR primer for KDR coding sequence.
XX
KW   KDR; VEGF receptor; vascular endothelial growth factor; inhibitor;
KW   neovascularisation suppressor; solid tumour; therapy; PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9831794-A1.
XX
PD   23-JUL-1998.
XX
PF   16-JAN-1998;   98WO-JP00140.
XX
PR   17-JAN-1997;   97JP-0019706.
XX
PA   (TOAG ) TOA GOSEI KK.
XX
PI   Matsumoto T,  Niwa M,  Okamoto M,  Segawa T;
XX
DR   WPI; 1998-414094/35.
XX
PT   Low-molecular vascular endothelial growth factor inhibiting peptide
PT   - for suppression of neovascularisation contains first and second
PT   immunoglobulin-like domains of KDR VEGF receptor
XX
PS   Example 3; Page 22; 53pp; Japanese.
XX
CC   This sequence is a PCR primer for DNA encoding the KDR protein, which is
CC   a vascular endothelial growth factor (VEGF) receptor. The polypeptide of
CC   the invention has a relatively low molecular weight and which inhibits
CC   the activity of VEGF. The polypeptide contains the first and second
CC   immunoglobulin-like (IGL) domains (and optionally the third) of the
CC   extracellular part of the KDR VEGF receptor but does not contain at least
CC   part of the sixth IGL domain. The peptide is used for the suppression of
CC   neovascularisation and thus treatment of diseases involving this, such as
CC   solid tumours. Because of its low molecular weight, the polypeptide
CC   readily infiltrates into the affected tissues.
XX
SQ   Sequence 42 BP; 12 A; 11 C; 8 G; 11 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 19;  Length 42;
  Best Local Similarity   84.2%;  Pred. No. 1.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAATC 19
           || ||||||||||  ||||
Db      29 TGAGTCCGGAGCCCAAATC 11


RESULT 4
AAH19643
```

```
ID   AAH19643 standard; DNA; 66 BP.
XX
AC   AAH19643;
XX
DT   20-JUL-2001  (first entry)
XX
DE   Human immunodeficiency virus gp120 reverse primer.
XX
KW   Human immunodeficiency virus; HIV; gp120; glycoprotein 120; antiviral;
KW   vaccine; CD4; chimeric polypeptide; FLSC; viral infection;
KW   PCR primer; ss.
XX
OS   Human immunodeficiency virus.
XX
PN   WO200127294-A1.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US27546.
XX
PR   08-OCT-1999;   99US-0158321.
XX
PA   (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.
XX
PI   Devico AL,  Fouts TR,  Tuskan RG;
XX
DR   WPI; 2001-290724/30.
XX
PT   Chimeric polypeptide comprising a virus coat polypeptide linked to a
PT   viral receptor polypeptide where the coat and receptor polypeptides can
PT   bind to each other, useful for preventing, inhibiting, or ameliorating
PT   a viral infection -
XX
PS   Example 1; Page 40; 80pp; English.
XX
CC   The present sequence was used in the construction of a polynucleotide
CC   encoding a single chain gp120-CD4 chimeric polypeptide FLSC. This
CC   procedure was described in an example illustrating an invention
CC   relating to a new chimeric polypeptide comprising a virus coat
CC   protein and a viral receptor protein linked by a spacer. The coat
CC   protein sequence and the viral receptor sequence interact with each
CC   other to form a complex that is capable of binding a co-receptor.
CC   The chimeric polypeptide is useful for preventing, inhibiting, or
CC   ameliorating a viral infection by passive protection against virus
CC   infection or by producing an immune response. Antibodies against the
CC   chimeric polypeptide and polynucleotides encoding the chimeric
CC   polypeptide are also useful for treating or preventing a viral
CC   infection. The chimeric polypeptide, polynucleotide and antibodies
CC   are also useful for diagnosing subjects at risk of a viral infection.
CC   The polypeptide is useful for identifying novel co-receptors and for
CC   characterising proteins as co-receptors.
XX
SQ   Sequence 66 BP; 10 A; 28 C; 20 G; 8 T; 0 other;

  Query Match            71.0%;  Score 14.2;  DB 22;  Length 66;
  Best Local Similarity  84.2%;  Pred. No. 1.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATCA 20
           ||||||||||||  ||  ||
Db       1 GGGTCCGGAGCCCGAGCCA 19


RESULT 5
AAL27594/c
ID   AAL27594 standard; DNA; 51 BP.
XX
AC   AAL27594;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #802.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1613; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 7 A; 18 C; 10 G; 16 T; 0 other;

  Query Match            69.0%;  Score 13.8;  DB 22;  Length 51;
  Best Local Similarity  88.2%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAA 17
           | |||||||||  |||||
Db      48 TTGGTCCGGAGGCTGAA 32


RESULT 6
ABN35552/c
ID   ABN35552 standard; DNA; 60 BP.
XX
AC   ABN35552;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:8300.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
```

OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 8300; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 12 A; 18 C; 12 G; 18 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 60;
Best Local Similarity 88.2%; Pred. No. 2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGGGTCCGGAGCCTGAA 17
|||||||||| || |||
Db 59 TGGGTCCGGATCCAGAA 43

RESULT 7
AAT74039
ID AAT74039 standard; cDNA; 24 BP.
XX
AC AAT74039;
XX
DT 16-SEP-1997 (first entry)
XX
DE Male-enhanced antigen-2 structural gene PCR primer.
XX
KW Mouse; MEA-2; detecting mutation; polymerase chain reaction; ss.
XX
OS Synthetic.
XX
PN JP09121869-A.
XX
PD 13-MAY-1997.
XX
PF 07-NOV-1995; 95JP-0311638.
XX
PR 07-NOV-1995; 95JP-0311638.
XX
PA (ITOH-) ITO HAM KK.
XX
DR WPI; 1997-314229/29.
XX
PT Male-enhanced antigen Mea-2 gene - especially from mouse, useful for
PT detecting mutation(s)
XX
PS Claim 9; Page 6; 13pp; Japanese.
XX
CC The present sequence represents a PCR primer corresponding to part
CC of the base sequence encoding male-enhanced antigen-2 (MEA-2). The
CC primer was used to amplify part of the structural gene of mouse MEA-2.
CC The polynucleotide encoding the whole MEA-2 sequence can be used for
CC the detection of mutations affecting the MEA-2 gene.
XX
SQ Sequence 24 BP; 5 A; 5 C; 9 G; 5 T; 0 other;

Query Match 68.0%; Score 13.6; DB 18; Length 24;
Best Local Similarity 80.0%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TGGGTCCGGAGCCTGAATCA 20
| || ||||| || ||||||
Db 1 TTGGGCCGGAACCGGAATCA 20

RESULT 8
AAX01737
ID AAX01737 standard; DNA; 28 BP.
XX
AC AAX01737;
XX
DT 07-JUN-1999 (first entry)
XX
DE Rat GDNFR-beta cDNA cloning upstream primer.
XX
KW Glial Cell Line-Derived Neurotrophic Factor; GDNF; GDNF receptor; GDNFR;
KW tyrosine phosphorylation; motorneuronal; neurodegenerative; human; rat;
KW Parkinson's disease; GDNFR-beta; PCR primer; ss.
XX
OS Synthetic.
XX
PN WO9852591-A1.
XX
PD 26-NOV-1998.
XX
PF 04-MAY-1998; 98WO-US09056.
XX
PR 22-MAY-1997; 97US-0861990.
XX
PA (CEPH-) CEPHALON INC.
XX
PI Arumae U, Ibanez CF, Saarma M, Sariola H, Suvanto P;
PI Trupp M;
XX
DR WPI; 1999-070126/06.
XX
PT Newly isolated and characterised GDNFR - used to further understand
PT its biological action and the signalling events that occur when
PT GDNFs bind to responsive cells
XX
PS Example 24; Page 54; 113pp; English.
XX
CC The invention relates to isolated Glial Cell Line-Derived Neurotrophic

```
CC   Factor Receptor (GDNFR) polypeptides. Isolated nucleic acid sequences
CC   encoding rat and human GDNFR-beta polypeptides are disclosed. GDNFR can
CC   be used in methods of identifying and isolating receptors. They can also
CC   be used for determining compounds/compositions that bind GDNF receptors.
CC   Similarly, the receptor is used to identify GDNF homologues by screening
CC   compounds/compositions exhibiting similar biological effects including
CC   tyrosine phosphorylation, and increases in c-fos mRNA and cell survival.
CC   GDNF analogues can be identified by screening compounds/compositions,
CC   which are antagonistic for the above biological effects. GDNF may have
CC   potential therapeutic use against motorneuronal and neurodegenerative
CC   diseases, especially Parkinson's disease. The isolation and
CC   characterisation of GDNF receptors will enable better understanding of
CC   its biological actions and the signalling events occurring when GDNF
CC   binds to responsive cells, which had previously been hindered by the
CC   lack of cell lines comprising GDNF receptors. Sequences AAX01737-38
CC   represent PCR primers used for cloning the cDNA encoding a rat
CC   GDNFR-beta polypeptide.
XX
SQ   Sequence 28 BP; 9 A; 8 C; 6 G; 5 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 20;  Length 28;
  Best Local Similarity   80.0%;  Pred. No. 2.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAATCA 20
           ||| |||| | ||||||| |
Db       2 TGGATCCGCAACCTGAATGA 21


RESULT 9
ABN55425
ID   ABN55425 standard; DNA; 65 BP.
XX
AC   ABN55425;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse spliced transcript detection oligonucleotide SEQ ID NO:28173.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Mus musculus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 28173; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 16 A; 12 C; 23 G; 14 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 65;
  Best Local Similarity   80.0%;  Pred. No. 2.6e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAATCA 20
           ||||||| |||   ||||||
Db       3 TGGGTCCAGAGTGGGAATCA 22


RESULT 10
ABN56524/c
ID   ABN56524 standard; DNA; 65 BP.
XX
AC   ABN56524;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse spliced transcript detection oligonucleotide SEQ ID NO:29272.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Mus musculus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 29272; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
```

```
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 21 A; 16 C; 13 G; 15 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 65;
  Best Local Similarity   80.0%;  Pred. No. 2.6e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAATCA 20
           | ||||||| | |||||| |
Db      42 TTGGTCCGGTGTCTGAATAA 23


RESULT 11
ABL57923/c
ID   ABL57923 standard; cDNA; 96 BP.
XX
AC   ABL57923;
XX
DT   22-JUL-2002  (first entry)
XX
DE   Human VG51 exon 4.
XX
KW   Antiasthmatic; anxiolytic; antiepileptic; antihypertensive; human;
KW   psychotropic; glutamate transporter; transporter; GABA;
KW   gamma-aminobutyric acid transporter; GABA transporter; neurotransmitter;
KW   asthma; anxiety; epilepsy; hypertension; psychiatric disorder;
KW   neurotic disorder; VG51; ss.
XX
OS   Homo sapiens.
XX
PN   WO200071709-A1.
XX
PD   30-NOV-2000.
XX
PF   19-MAY-2000; 2000WO-FR01383.
XX
PR   21-MAY-1999;   99FR-0006525.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Giros B,  Gasnier B,  Sagne C,  El Mestikawy S,  Hamon M;
XX
DR   WPI; 2001-025160/03.
XX
PT   New mammalian amino acid transporter, used e.g. to screen for
PT   psychotropic agents, is high capacity but low affinity transporter of
PT   gamma-aminobutyric acid  -
XX
PS   Claim 5; Page 96; 103pp; French.
XX
CC   The present sequence is exon 4 of human VG51 cDNA sequence, a glutamate/
CC   gamma-aminobutyric acid (GABA) transporter. GABA and glutamate are
CC   neurotransmitters. The transporter can be used to produce specific
CC   antibodies, to screen for binding agents. Modulators of the transporter
CC   are useful for treating disorders associated with deregulated
CC   glutamate/GABA transport, e.g. asthma, anxiety, epilepsy, hypertension
CC   and other psychiatric and neurotic disorders, while determining levels of
CC   the transporter and its coding sequence can be used for diagnosis of such
CC   disorders. The full length coding sequence is given in ABL57930.
XX
SQ   Sequence 96 BP; 23 A; 23 C; 26 G; 24 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 23;  Length 96;
  Best Local Similarity   80.0%;  Pred. No. 2.6e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAATCA 20
           ||| |||||||| | | ||
Db      81 TGGTTCCGGAGCCAGGAGCA 62


RESULT 12
AAA94347
ID   AAA94347 standard; DNA; 21 BP.
XX
AC   AAA94347;
XX
DT   10-JAN-2001  (first entry)
XX
DE   Oligonucleotide 1c for bacterial species identification.
XX
KW   Large ribosomal subunit; 23S rDNA; bacterial species identification;
KW   oligonucleotide probe; ss.
XX
OS   Escherichia coli.
XX
PN   WO200052203-A2.
XX
PD   08-SEP-2000.
XX
PF   01-MAR-2000; 2000WO-GB00740.
XX
PR   02-MAR-1999;   99GB-0004804.
XX
PA   (UNLO ) KINGS COLLEGE LONDON.
PA   (GUYS-) GUY'S & ST THOMAS'S NAT HEALTH TRUST.
XX
PI   French GL,  Anthony RM,  Brown TJ;
XX
DR   WPI; 2000-587320/55.
XX
PT   Detecting bacteria, e.g. Escherichia coli, in a sample, by amplifying
PT   portion of 23S rDNA using primers of defined sequences and testing the
PT   resulting amplicon by reverse hybridization with oligonucleotide probes
PT   -
XX
PS   Claim 6; Page 5; 39pp; English.
XX
CC   The present sequence is one of a number of oligonucleotide used in a
CC   method for identifying bacteria in a sample. PCR primers were used to
CC   amplify the large ribosomal subunit (23S rDNA) from bacteria in the
CC   sample. The resulting amplicon was then used to probe oligonucleotides,
CC   including the present sequence, which were designed to hybridise to only
CC   one bacterial species. This method is useful for detecting at least one
CC   of Escherichia coli, Staphylococcus aureus, Pseudomonas aeruginosa,
CC   Enterococcus spp. preferably Enterococcus feacium or Enterococcus
CC   feacalis, Klebsiella spp., Enterobacter spp., Proteus spp. preferably
CC   Proteus mirabilis, Pneumococci, coagulase negative Staphylococci,
CC   Listeria sp., Stenorophomonas maltophilia and Burkholderia cepacia. The
CC   oligonucleotide probes can be used singly for the identification of
CC   certain individual species or in a panel or array for the identification
CC   of many different species. The assay can be easily extended to identify a
CC   wider range of bacterial species with the addition of oligonucleotides
CC   without increasing the complexity of performing the assay. Unlike prior
CC   art techniques, the detection method rapidly identifies large numbers of
CC   bacteria without prior knowledge of the causative agents.
XX
SQ   Sequence 21 BP; 5 A; 5 C; 6 G; 5 T; 0 other;
```

```
  Query Match             67.0%;  Score 13.4;  DB 21;  Length 21;
  Best Local Similarity   93.3%;  Pred. No. 3e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 CCGGAGCCTGAATCA 20
          || ||||||||||||
Db      1 CCAGAGCCTGAATCA 15


RESULT 13
AAF28284/c
ID   AAF28284 standard; DNA; 23 BP.
XX
AC   AAF28284;
XX
DT   04-APR-2001  (first entry)
XX
DE   Oligonucleotide SPF.
XX
KW   Glucocerebrosidase; Gcc; splice; gene therapy; enzyme therapy;
KW   Gaucher; ss.
XX
OS   Unidentified.
XX
PN   CA2272055-A1.
XX
PD   02-DEC-2000.
XX
PF   02-JUN-1999;   99CA-2272055.
XX
PR   02-JUN-1999;   99CA-2272055.
XX
PA   (HSCR-) HSC RES & DEV LP.
XX
PI   Mahuran DJ,  Callahan JW,  Clarke JTR;
XX
DR   WPI; 2001-160192/17.
XX
PT   New glucocerebrosidase (Gcc) DNA molecule modified in at least one
PT   nucleotide, useful in gene therapy or enzyme replacement therapy,
PT   particularly in treating Gaucher disease  -
XX
PS   Disclosure; Page 21; 37pp; English.
XX
CC   The present invention relates to glucocerebrosidase (Gcc) DNA
CC   with a modification in at least one nucleotide that disrupts a
CC   splicing consensus sequence and prevents splicing of mRNA produced
CC   from the DNA molecule, while preserving the ability of the DNA to
CC   express active Gcc. The glucocerebrosidase (Gcc) DNA molecule and
CC   the vector comprising the DNA are useful in gene therapy or enzyme
CC   replacement therapy, particularly in treating Gaucher disease. The
CC   cells transfected or transduced by the vectors are useful as research
CC   tools.
XX
SQ   Sequence 23 BP; 4 A; 10 C; 5 G; 4 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 22;  Length 23;
  Best Local Similarity   83.3%;  Pred. No. 3.8e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGGTCCGGAGCCTGAATC 19
          | |||||||| ||| |||
Db     22 GAGTCCGGAGGCTGGATC 5


RESULT 14
AAS15354/c
ID   AAS15354 standard; DNA; 23 BP.
XX
AC   AAS15354;
XX
DT   13-FEB-2002  (first entry)
XX
DE   pCMV/SPORT vector primer 2182-36.
XX
KW   Interleukin 17; hIL-17 receptor like protein; immunomodulatory;
KW   anti-inflammatory; anti-diabetic; immunosuppressive; anti-microbial;
KW   hepatic; anabolic; anorectic; anti-alzheimer's; anti-parkinsonian;
KW   anti-convulsant; anti-asthmatic; dermatological; renal; osteopathic;
KW   vascular; cytostatic; anti-leukaemic; anti-infertility; ophthalmological;
KW   hepatitis; anorexia; cachexia; neuronal dysfunction; lung disease;
KW   bone disease; vascular disorder; eye disorder; cancer; PCR primer;
KW   human; ss.
XX
OS   Synthetic.
XX
PN   WO200168859-A2.
XX
PD   20-SEP-2001.
XX
PF   15-MAR-2001; 2001WO-US08678.
XX
PR   16-MAR-2000; 2000US-189816P.
PR   28-NOV-2000; 2000US-0724460.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Jing S;
XX
DR   WPI; 2001-611392/70.
XX
PT   Nucleic acids encoding interleukin 17 receptor like polypeptides,
PT   useful for preventing, diagnosing and treating, e.g. leukaemia, asthma,
PT   diabetes, psoriasis and glaucoma -
XX
PS   Example 1; Page 124; 158pp; English.
XX
CC   The invention describes novel nucleic acids encoding interleukin (IL) 17
CC   receptor like polypeptides useful as vaccines and in gene therapy. These
CC   have immunomodulatory, anti-inflammatory, anti-diabetic, anti-microbial,
CC   immunosuppressive, hepatic, anabolic, anorectic, anti-alzheimer's, renal,
CC   anti-parkinsonian, anti-convulsant, anti-asthmatic, dermatological,
CC   osteopathic, vascular, cytostatic, anti-leukaemic, anti-infertility and
CC   ophthalmological activities. The IL-17 receptor like nucleic acids and
CC   proteins may be used to prevent and treat diseases associated with
CC   inappropriate IL-17 receptor like polypeptide (IL17rlp) expression. These
CC   include, for example immune disorders (e.g. inflammation, diabetes and
CC   transplant rejection), infections (e.g. hepatitis and septicaemia),
CC   weight disorders (e.g. anorexia, cachexia and obesity), neuronal
CC   dysfunction (e.g. Alzheimer's disease, Parkinson's disease and epilepsy),
CC   lung disorders (e.g. cystic fibrosis, asthma and emphysema), skin disease
CC   (e.g. eczema and psoriasis), kidney disease (e.g. glomerulonephritis),
CC   bone diseases (e.g. osteoporosis and hypercalcaemia), vascular disorders
CC   (e.g. stroke and atherosclerosis, cancers (e.g. leukaemia, myeloma and
CC   breast cancer), reproductive disorders (e.g. infertility and
CC   miscarriage), eye disorders (e.g. glaucoma and retinal neuropathy). The
CC   DNA and its complements may also used as diagnostic probes to detect and
CC   quantitate the presence of similar nucleic acids in samples and identify
CC   patients needing restorative therapy. The IL17rlp may also be used as
CC   antigens in the production of antibodies against the proteins and in
CC   assays to identify modulators of expression and activity. The
CC   anti-IL17rlp antibodies and antagonists may also be used to down regulate
CC   expression and activity. This sequence is the pCMV/SPORT vector PCR
CC   primer (2182-36), specific to Interleukin 17 (IL-17) receptor like
CC   protein and used with PCR primer AAS15353 to isolate the gene from a
CC   human tissue library described in the method of the invention.
XX
SQ   Sequence 23 BP; 4 A; 10 C; 5 G; 4 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 23;  Length 23;
  Best Local Similarity   83.3%;  Pred. No. 3.8e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGGTCCGGAGCCTGAATC 19
          | |||||||| ||| |||
Db     20 GAGTCCGGAGGCTGGATC 3
```

```
RESULT 15
AAZ69457
ID   AAZ69457 standard; DNA; 47 BP.
XX
AC   AAZ69457;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:3813.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation;
KW   diagnosis; single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,G)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB00822.
XX
PR   21-APR-1998;   98US-0082614.
PR   23-NOV-1998;   98US-0109732.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome  -
XX
PS   Claim 3; Page 1044; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the
CC   invention have a variety of uses: they can be used for high density
CC   mapping of the human genome, and in complex association studies and
CC   haplotyping studies which are useful in determining the genetic basis
CC   for disease states. Compositions and methods of the invention can also
CC   be useful for the identification of the targets for the development of
CC   pharmaceutical agents and diagnostic methods, as well as the
CC   characterisation of the differential efficacious responses to and side
CC   effects from pharmaceutical agents acting on a disease as well as other
CC   treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC   and 3367, are not actually given a sequence in the Sequence Listing
CC   from the present invention.
XX
SQ   Sequence 47 BP; 17 A; 10 C; 11 G; 9 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 21;  Length 47;
  Best Local Similarity   83.3%;  Pred. No. 4e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GGTCCGGAGCCTGAATCA 20
           ||||| ||| ||||||  |
Db      19 GGTCCAGAGACTGAATGA 36


Search completed: November 23, 2002, 06:28:38
```

Job time : 102.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-4
Perfect score: 20
Sequence: 1 tgggtccggagcctgaatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	42	4	US-09-142-956B-8	Sequence 8, Appli
	2	13.8	69.0	20	1	US-08-586-274-22	Sequence 22, Appl
	3	12.8	64.0	18	3	US-09-357-072-8	Sequence 8, Appli
	4	12.8	64.0	38	1	US-07-744-282C-47	Sequence 47, Appl
	5	12.8	64.0	38	5	PCT-US92-06821A-90	Sequence 90, Appl
c	6	12.8	64.0	47	4	US-09-641-638-691	Sequence 691, App
	7	12.8	64.0	50	2	US-08-244-548-5	Sequence 5, Appli
c	8	12.8	64.0	50	2	US-08-244-548-6	Sequence 6, Appli
c	9	12.6	63.0	19	4	US-09-262-773-33	Sequence 33, Appl
	10	12.6	63.0	49	4	US-09-538-709-1064	Sequence 1064, Ap
c	11	12.6	63.0	69	2	US-08-599-455B-20	Sequence 20, Appl
c	12	12.6	63.0	69	4	US-09-069-781B-20	Sequence 20, Appl
c	13	12.6	63.0	69	4	US-09-137-132-20	Sequence 20, Appl
c	14	12.6	63.0	69	4	US-08-864-564A-20	Sequence 20, Appl
c	15	12.6	63.0	69	4	US-09-094-410-20	Sequence 20, Appl
	16	12.6	63.0	97	4	US-09-020-956-64	Sequence 64, Appl
	17	12.6	63.0	97	4	US-09-030-607-64	Sequence 64, Appl
	18	12.6	63.0	97	4	US-09-605-785-64	Sequence 64, Appl
	19	12.6	63.0	97	4	US-09-439-313-64	Sequence 64, Appl
	20	12.6	63.0	97	4	US-09-352-616A-64	Sequence 64, Appl
	21	12.6	63.0	97	4	US-09-232-149A-64	Sequence 64, Appl
	22	12.4	62.0	24	5	PCT-US93-02457-11	Sequence 11, Appl
	23	12.4	62.0	24	5	PCT-US93-11458-12	Sequence 12, Appl
	24	12.4	62.0	28	5	PCT-US93-11458-13	Sequence 13, Appl
c	25	12.2	61.0	20	2	US-09-166-203-50	Sequence 50, Appl
c	26	12.2	61.0	20	4	US-09-377-309-50	Sequence 50, Appl
	27	12.2	61.0	24	4	US-09-600-031-9	Sequence 9, Appli
c	28	12.2	61.0	28	4	US-08-460-269C-15	Sequence 15, Appl
c	29	12.2	61.0	36	1	US-08-244-492A-9	Sequence 9, Appli
c	30	12.2	61.0	36	1	US-08-709-915-12	Sequence 12, Appl
	31	12.2	61.0	60	4	US-08-264-578-4	Sequence 4, Appli
	32	12.2	61.0	60	4	US-08-264-578-5	Sequence 5, Appli
	33	12.2	61.0	65	1	US-08-484-686B-34	Sequence 34, Appl
	34	12.2	61.0	65	1	US-08-484-686B-37	Sequence 37, Appl
	35	12.2	61.0	65	4	US-08-463-160B-34	Sequence 34, Appl
	36	12.2	61.0	65	4	US-08-463-160B-37	Sequence 37, Appl
	37	12.2	61.0	65	5	PCT-US91-02568-38	Sequence 38, Appl
	38	12.2	61.0	68	2	US-08-790-963-87	Sequence 87, Appl
	39	12.2	61.0	68	4	US-09-371-774-87	Sequence 87, Appl
c	40	12.2	61.0	69	3	US-08-812-121-13	Sequence 13, Appl
c	41	12.2	61.0	69	4	US-09-403-672-13	Sequence 13, Appl
	42	12	60.0	26	4	US-08-463-691-17	Sequence 17, Appl
	43	12	60.0	26	4	US-08-255-236-5	Sequence 5, Appli
	44	12	60.0	30	3	US-08-961-083-327	Sequence 327, App
	45	12	60.0	78	4	US-09-357-487B-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-142-956B-8/c
; Sequence 8, Application US/09142956B
; Patent No. 6348333
; GENERAL INFORMATION:
; APPLICANT: Niwa, Mikio
; APPLICANT: Okamoto, Masaji
; APPLICANT: Matsumoto, Tomoe
; APPLICANT: Segawa, Toshiaki
; TITLE OF INVENTION: VEGF-BINDING POLYPEPTIDES
; FILE REFERENCE: 06501-021001
; CURRENT APPLICATION NUMBER: US/09/142,956B
; CURRENT FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: PCT/JP98/00140
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: JP 9/19706
; PRIOR FILING DATE: 1997-01-17
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated primer
US-09-142-956B-8

Query Match 71.0%; Score 14.2; DB 4; Length 42;
Best Local Similarity 84.2%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TGGGTCCGGAGCCTGAATC 19
           || |||||||||  ||||
Db      29 TGAGTCCGGAGCCCAAATC 11
```

RESULT 2
US-08-586-274-22
; Sequence 22, Application US/08586274
; Patent No. 5702895
; GENERAL INFORMATION:
; APPLICANT: MATSUNAGA, HIRONARI
; APPLICANT: TSUKUMO, KENICHI
; APPLICANT: WAKISAKA, SHINJI
; APPLICANT: YAMANE, AKIO
; TITLE OF INVENTION: METHOD AND KIT FOR DETECTING
; TITLE OF INVENTION: METHICILLIN-RESISTANT STAPHYLOCOCCUS AUREUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

```
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/586,274
;      FILING DATE:  16-JAN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-006390
;      FILING DATE:  19-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  209-038-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
US-08-586-274-22

  Query Match             69.0%;  Score 13.8;  DB 1;  Length 20;
  Best Local Similarity   88.2%;  Pred. No. 2.3e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GTCCGGAGCCTGAATCA 20
           ||||| | |||||||||
Db       2 GTCCGTAACCTGAATCA 18


RESULT 3
US-09-357-072-8
; Sequence 8, Application US/09357072
; Patent No. 6015712
; GENERAL INFORMATION:
;  APPLICANT: Brett P. Monia
;  APPLICANT: Brenda F. Baker
;  APPLICANT: Hong Zhang
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF FADD EXPRESSION
;  FILE REFERENCE: RTS-0027
;  CURRENT APPLICATION NUMBER: US/09/357,072
;  CURRENT FILING DATE: 1999-07-19
;  NUMBER OF SEQ ID NOS: 87
; SEQ ID NO 8
;   LENGTH: 18
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense Oligonucleotide
US-09-357-072-8

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 18;
  Best Local Similarity   87.5%;  Pred. No. 7.3e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 TCCGGAGCCTGAATCA 20
           ||||| |||||| |||
Db       2 TCCGGTGCCTGATTCA 17


RESULT 4
US-07-744-282C-47
; Sequence 47, Application US/07744282C
; Patent No. 5521300
;  GENERAL INFORMATION:
;    APPLICANT:  Shah, Jyotsna S.
;    APPLICANT:  Nietupski, Raymond M.
;    APPLICANT:  Liu, Jing
;    TITLE OF INVENTION:  Oligonucleotides Complementary to
;    TITLE OF INVENTION:  Mycobacterial Nucleic Acids
;    NUMBER OF SEQUENCES:  127
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kevin M. Farrell, P.C.
;      STREET:  P.O. Box 999
;      CITY:  York Harbor
;      STATE:  ME
;      COUNTRY:  USA
;      ZIP:  03911
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/744,282C
;      FILING DATE:  August 13, 1991
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kevin M. Farrell
;      REGISTRATION NUMBER:  35,505
;      REFERENCE/DOCKET NUMBER:  GTR90-05
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (207) 363-0558
;      TELEFAX:  (207) 363-0528
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA (genomic)
US-07-744-282C-47

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 38;
  Best Local Similarity   75.0%;  Pred. No. 7.7e+02;
  Matches   12;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GTCCGGAGCCTGAATC 19
           | || |||||:|||:|
Db      23 GCCCAGAGCCUGAAUC 38


RESULT 5
PCT-US92-06821A-90
; Sequence 90, Application PC/TUS9206821A
;  GENERAL INFORMATION:
;    APPLICANT:  Shah, Jyotsna S.
;    APPLICANT:  Nietupski, Raymond M.
;    APPLICANT:  Liu, Jing
;    TITLE OF INVENTION:  Oligonucleotides Complementary to
;    TITLE OF INVENTION:  Mycobacterial Nucleic Acids
;    NUMBER OF SEQUENCES:  133
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amoco Corporation
;      STREET:  200 East Randolph Drive, P.O. Box 87703
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  U.S.A.
;      ZIP:  60680
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06821A
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/744,282
;      FILING DATE:  13-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Galloway, Norval B.
;      REGISTRATION NUMBER:  33,595
;      REFERENCE/DOCKET NUMBER:  CN 5851
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-856-7180
;      TELEFAX:  312-856-4972
;  INFORMATION FOR SEQ ID NO:  90:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
PCT-US92-06821A-90

  Query Match             64.0%;  Score 12.8;  DB 5;  Length 38;
  Best Local Similarity   75.0%;  Pred. No. 7.7e+02;
  Matches   12;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GTCCGGAGCCTGAATC 19
           | || |||||:|||:|
Db      23 GCCCAGAGCCUGAAUC 38


RESULT 6
US-09-641-638-691/c
; Sequence 691, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Cohen, Annick
;  TITLE OF INVENTION:  BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
;  TITLE OF INVENTION:  GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
;  FILE REFERENCE:  GENSET.051CP1
;  CURRENT APPLICATION NUMBER: US/09/641,638
;  CURRENT FILING DATE:  2000-08-16
;  PRIOR APPLICATION NUMBER:  US 09/502,330
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 691
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 10-349-97 : polymorphic base A or G
US-09-641-638-691

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 47;
  Best Local Similarity   77.8%;  Pred. No. 7.9e+02;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATC 19
           |||:  |||| |||||||
Db      27 GGGYTGGGAGGCTGAATC 10


RESULT 7
US-08-244-548-5
; Sequence 5, Application US/08244548
; Patent No. 5874556
;  GENERAL INFORMATION:
;    APPLICANT:  LUPTON, STEPHEN D.
;    APPLICANT:  ALLEN, JAMES M.
;    TITLE OF INVENTION:  HYBRID GENES FOR USE IN THE PRODUCTION
;    TITLE OF INVENTION:  OF TH-INDEPENDENT CYTOTOXIC T CELLS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/244,548
;      FILING DATE:  06-JUN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DYLAN, TYLER M.
;      REGISTRATION NUMBER:  37,612
;      REFERENCE/DOCKET NUMBER:  22627-20005.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-244-548-5

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 50;
  Best Local Similarity   87.5%;  Pred. No. 7.9e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GTCCGGAGCCTGAATC 19
           ||| ||||| ||||||
Db      23 GTCAGGAGCTTGAATC 38


RESULT 8
US-08-244-548-6/c
; Sequence 6, Application US/08244548
; Patent No. 5874556
;  GENERAL INFORMATION:
;    APPLICANT:  LUPTON, STEPHEN D.
;    APPLICANT:  ALLEN, JAMES M.
;    TITLE OF INVENTION:  HYBRID GENES FOR USE IN THE PRODUCTION
;    TITLE OF INVENTION:  OF TH-INDEPENDENT CYTOTOXIC T CELLS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/244,548
;      FILING DATE:  06-JUN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DYLAN, TYLER M.
;      REGISTRATION NUMBER:  37,612
;      REFERENCE/DOCKET NUMBER:  22627-20005.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-244-548-6

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 50;
  Best Local Similarity   87.5%;  Pred. No. 7.9e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GTCCGGAGCCTGAATC 19
           ||| ||||| ||||||
Db      32 GTCAGGAGCTTGAATC 17


RESULT 9
US-09-262-773-33/c
; Sequence 33, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
;  APPLICANT: Ballinger, Dennis G.
;  APPLICANT: Ding, Wei
;  APPLICANT: Wagner, Susanne
;  APPLICANT: Hess, Mark A.
;  TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
;  FILE REFERENCE: Myriad 3
;  CURRENT APPLICATION NUMBER: US/09/262,773
;  CURRENT FILING DATE: 1999-03-04
;  NUMBER OF SEQ ID NOS: 210
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
;   LENGTH: 19
;   TYPE: DNA
;   ORGANISM: primer
US-09-262-773-33

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 19;
  Best Local Similarity   78.9%;  Pred. No. 9.2e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAATC 19
           | ||  |||||||||| ||
Db      19 TAGGAGCGGAGCCTGAGTC 1


RESULT 10
US-09-538-709-1064
; Sequence 1064, Application US/09538709
; Patent No. 6468749
; GENERAL INFORMATION:
;  APPLICANT: Ulanovsky, et al
;  TITLE OF INVENTION: SEQUENCE-DEPENDENT GENE SORTING TECHNIQUES
;  FILE REFERENCE: 540579-2006
;  CURRENT APPLICATION NUMBER: US/09/538,709
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 1311
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1064
;   LENGTH: 49
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Adaptor
US-09-538-709-1064

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 49;
  Best Local Similarity   78.9%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATCA 20
           ||  |||||||||| |||
Db      23 GGCCGCGGAGCCTGATTCA 41


RESULT 11
US-08-599-455B-20/c
; Sequence 20, Application US/08599455B
; Patent No. 5972621
;  GENERAL INFORMATION:
;    APPLICANT:  Tartaglia, Louis A.
;    APPLICANT:  Tepper, Robert I.
;    APPLICANT:  Culpepper, Janice A.
;    TITLE OF INVENTION:  METHODS OF IDENTIFYING COMPOUNDS THAT
;    TITLE OF INVENTION:  MODULATE BODY WEIGHT USING THE OB RECEPTOR
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/599,455B
;      FILING DATE:  22-JAN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/583,153
;      FILING DATE:  28-DEC-1995
;      APPLICATION NUMBER:  08/570,142
;      FILING DATE:  11-DEC-1995
;      APPLICATION NUMBER:  08/569,485
;      FILING DATE:  08-DEC-1995
;      APPLICATION NUMBER:  08/566,622
;      FILING DATE:  04-DEC-1995
;      APPLICATION NUMBER:  08/562,663
;      FILING DATE:  27-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meiklejohn, Ph.D., Anita L.
;      REGISTRATION NUMBER:  35,283
;      REFERENCE/DOCKET NUMBER:  07334/017001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
US-08-599-455B-20

  Query Match             63.0%;  Score 12.6;  DB 2;  Length 69;
```

Best Local Similarity 78.9%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       2 GGGTCCGGAGCCTGAATCA 20
           || |  |||||||||| ||
Db      25 GGCTTTGGAGCCTGAAACA 7
```

```
RESULT 12
US-09-069-781B-20/c
; Sequence 20, Application US/09069781B
; Patent No. 6287782
;  GENERAL INFORMATION:
;    APPLICANT:  Tartaglia, Louis A.
;    APPLICANT:  Tepper, Robert I.
;    APPLICANT:  Culpepper, Janice A.
;    APPLICANT:  White, David W.
;    TITLE OF INVENTION:  THE OB RECEPTOR AND METHODS FOR
;    TITLE OF INVENTION:  THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
;    TITLE OF INVENTION:  INCLUDING OBESITY AND CACHEXIA
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/069,781B
;      FILING DATE:  29-APRIL-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/864,564
;      FILING DATE:  28-MAY-1997
;      APPLICATION NUMBER:  US 08/708,123
;      FILING DATE:  03-SEP-1996
;      APPLICATION NUMBER:  US 08/638,524
;      FILING DATE:  26-APR-1996
;      APPLICATION NUMBER:  US 08/599,455
;      FILING DATE:  22-JAN-1996
;      APPLICATION NUMBER:  US 08/583,153
;      FILING DATE:  28-DEC-1995
;      APPLICATION NUMBER:  US 08/570,142
;      FILING DATE:  11-DEC-1995
;      APPLICATION NUMBER:  US 08/569,485
;      FILING DATE:  08-DEC-1995
;      APPLICATION NUMBER:  US 08/566,622
;      FILING DATE:  04-DEC-1995
;      APPLICATION NUMBER:  US 08/562,663
;      FILING DATE:  27-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meiklejohn, Ph.D., Anita L.
;      REGISTRATION NUMBER:  35,283
;      REFERENCE/DOCKET NUMBER:  07334/082001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  RNA
US-09-069-781B-20
```

Query Match 63.0%; Score 12.6; DB 4; Length 69;

Best Local Similarity 78.9%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GGGTCCGGAGCCTGAATCA 20
|| | |||||||||| ||
Db 25 GGCTTTGGAGCCTGAAACA 7

```
RESULT 13
US-09-137-132-20/c
; Sequence 20, Application US/09137132
; Patent No. 6380363
;  GENERAL INFORMATION:
;    APPLICANT:  Tartaglia, Louis A.
;    APPLICANT:  Tepper, Robert I.
;    APPLICANT:  Culpepper, Janice A.
;    APPLICANT:  White, David W.
;    TITLE OF INVENTION:  THE OB RECEPTOR AND METHODS FOR
;    TITLE OF INVENTION:  THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
;    TITLE OF INVENTION:  INCLUDING OBESITY AND CACHEXIA
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/137,132
;      FILING DATE:  18-AUG-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/864,564
;      FILING DATE:  28-MAY-1997
;      APPLICATION NUMBER:  08/708,123
;      FILING DATE:  03-SEP-1996
;      APPLICATION NUMBER:  08/638,524
;      FILING DATE:  26-APR-1996
;      APPLICATION NUMBER:  08/599,455
;      FILING DATE:  22-JAN-1996
;      APPLICATION NUMBER:  08/583,153
;      FILING DATE:  28-DEC-1995
;      APPLICATION NUMBER:  08/570,142
;      FILING DATE:  11-DEC-1995
;      APPLICATION NUMBER:  08/569,485
;      FILING DATE:  08-DEC-1995
;      APPLICATION NUMBER:  08/566,622
;      FILING DATE:  04-DEC-1995
;      APPLICATION NUMBER:  08/562,663
;      FILING DATE:  27-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meiklejohn, Ph.D., Anita L.
;      REGISTRATION NUMBER:  35,283
;      REFERENCE/DOCKET NUMBER:  07334/019004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  RNA
US-09-137-132-20
```

Query Match 63.0%; Score 12.6; DB 4; Length 69;

```
  Best Local Similarity   78.9%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATCA 20
           || |  |||||||||| ||
Db      25 GGCTTTGGAGCCTGAAACA 7


RESULT 14
US-08-864-564A-20/c
; Sequence 20, Application US/08864564A
; Patent No. 6395498
;  GENERAL INFORMATION:
;    APPLICANT:  Tartaglia, Louis A.
;    APPLICANT:  Tepper, Robert I.
;    APPLICANT:  Culpepper, Janice A.
;    APPLICANT:  White, David W.
;    TITLE OF INVENTION:  THE OB RECEPTOR AND METHODS FOR
;    TITLE OF INVENTION:  THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
;    TITLE OF INVENTION:  INCLUDING OBESITY AND CACHEXIA
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/864,564A
;      FILING DATE:  28-MAY-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/708,123
;      FILING DATE:  03-SEP-1996
;      APPLICATION NUMBER:  08/638,524
;      FILING DATE:  26-APR-1996
;      APPLICATION NUMBER:  08/599,455
;      FILING DATE:  22-JAN-1996
;      APPLICATION NUMBER:  08/583,153
;      FILING DATE:  28-DEC-1995
;      APPLICATION NUMBER:  08/570,142
;      FILING DATE:  11-DEC-1995
;      APPLICATION NUMBER:  08/569,485
;      FILING DATE:  08-DEC-1995
;      APPLICATION NUMBER:  08/566,622
;      FILING DATE:  04-DEC-1995
;      APPLICATION NUMBER:  08/562,663
;      FILING DATE:  27-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meiklejohn, Ph.D., Anita L.
;      REGISTRATION NUMBER:  35,283
;      REFERENCE/DOCKET NUMBER:  07334/019002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  RNA
US-08-864-564A-20

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 69;
  Best Local Similarity   78.9%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATCA 20
           || |  |||||||||| ||
Db      25 GGCTTTGGAGCCTGAAACA 7


RESULT 15
US-09-094-410-20/c
; Sequence 20, Application US/09094410
; Patent No. 6403552
;  GENERAL INFORMATION:
;    APPLICANT:  Tartaglia, Louis A.
;    APPLICANT:  Tepper, Robert I.
;    APPLICANT:  Culpepper, Janice A.
;    APPLICANT:  White, David W.
;    TITLE OF INVENTION:  THE OB RECEPTOR AND METHODS FOR
;    TITLE OF INVENTION:  THE DIAGNOSIS AND TREATMENT OF BODY WEIGHT DISORDERS,
;    TITLE OF INVENTION:  INCLUDING OBESITY AND CACHEXIA
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/094,410
;      FILING DATE:  09-JUN-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/864,564
;      FILING DATE:  28-MAY-1997
;      APPLICATION NUMBER:  08/708,123
;      FILING DATE:  03-SEP-1996
;      APPLICATION NUMBER:  08/638,524
;      FILING DATE:  26-APR-1996
;      APPLICATION NUMBER:  08/599,455
;      FILING DATE:  22-JAN-1996
;      APPLICATION NUMBER:  08/583,153
;      FILING DATE:  28-DEC-1995
;      APPLICATION NUMBER:  08/570,142
;      FILING DATE:  11-DEC-1995
;      APPLICATION NUMBER:  08/569,485
;      FILING DATE:  08-DEC-1995
;      APPLICATION NUMBER:  08/566,622
;      FILING DATE:  04-DEC-1995
;      APPLICATION NUMBER:  08/562,663
;      FILING DATE:  27-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meiklejohn, Ph.D., Anita L.
;      REGISTRATION NUMBER:  35,283
;      REFERENCE/DOCKET NUMBER:  07334/019003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  RNA
US-09-094-410-20

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 69;
  Best Local Similarity   78.9%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy       2 GGGTCCGGAGCCTGAATCA 20
           || |  |||||||||| ||
Db      25 GGCTTTGGAGCCTGAAACA 7
```

Search completed: November 23, 2002, 06:36:01
Job time : 23.55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-4
Perfect score: 20
Sequence: 1 tgggtccggagcctgaatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.2	71.0	66	9	US-09-934-060A-8	Sequence 8, Appli
c	2	13.2	66.0	23	10	US-09-809-567-11	Sequence 11, Appl
c	3	13.2	66.0	82	10	US-09-864-761-20574	Sequence 20574, A
	4	12.8	64.0	25	10	US-09-886-055-519	Sequence 519, App
	5	12.8	64.0	25	10	US-09-886-055-521	Sequence 521, App
c	6	12.6	63.0	33	10	US-09-780-668A-25	Sequence 25, Appl
	7	12.6	63.0	33	10	US-09-780-668A-26	Sequence 26, Appl
c	8	12.6	63.0	41	10	US-09-821-821-17	Sequence 17, Appl
	9	12.6	63.0	97	10	US-09-759-143-64	Sequence 64, Appl
	10	12.6	63.0	97	10	US-09-780-669-64	Sequence 64, Appl
	11	12.6	63.0	97	10	US-09-030-606-64	Sequence 64, Appl
	12	12.6	63.0	97	10	US-09-822-827-64	Sequence 64, Appl
	13	12.6	63.0	97	10	US-09-115-453-64	Sequence 64, Appl
	14	12.4	62.0	39	12	US-10-081-281-16	Sequence 16, Appl
c	15	12.4	62.0	66	10	US-09-969-708-357	Sequence 357, App
	16	12.4	62.0	69	10	US-09-775-743A-10	Sequence 10, Appl
c	17	12.4	62.0	81	8	US-08-790-540A-17	Sequence 17, Appl
c	18	12.4	62.0	81	8	US-08-791-391A-17	Sequence 17, Appl
c	19	12.4	62.0	81	8	US-08-791-391A-26	Sequence 26, Appl
	20	12.2	61.0	21	10	US-09-726-774-64	Sequence 64, Appl
c	21	12.2	61.0	32	10	US-09-814-777A-51	Sequence 51, Appl
c	22	12	60.0	18	10	US-09-809-867-1	Sequence 1, Appli
	23	12	60.0	30	10	US-09-765-272-327	Sequence 327, App
	24	12	60.0	54	10	US-09-051-034A-5	Sequence 5, Appli
	25	12	60.0	98	10	US-09-864-761-27818	Sequence 27818, A
c	26	12	60.0	98	10	US-09-969-708-151	Sequence 151, App
c	27	11.8	59.0	18	10	US-09-780-668A-22	Sequence 22, Appl
	28	11.8	59.0	18	10	US-09-780-668A-23	Sequence 23, Appl
	29	11.8	59.0	31	10	US-09-801-274-570	Sequence 570, App
c	30	11.8	59.0	82	10	US-09-908-855-3	Sequence 3, Appli
c	31	11.6	58.0	20	10	US-09-733-294A-107	Sequence 107, App
c	32	11.6	58.0	24	9	US-09-978-295A-99	Sequence 99, Appl
c	33	11.6	58.0	24	9	US-09-978-697-99	Sequence 99, Appl
c	34	11.6	58.0	27	10	US-09-759-143-832	Sequence 832, App
c	35	11.6	58.0	27	10	US-09-780-669-832	Sequence 832, App
c	36	11.6	58.0	27	10	US-09-822-827-832	Sequence 832, App
c	37	11.6	58.0	30	10	US-09-872-820-6	Sequence 6, Appli
c	38	11.6	58.0	54	12	US-10-113-573-1	Sequence 1, Appli
c	39	11.6	58.0	82	10	US-09-923-876-3530	Sequence 3530, Ap
	40	11.6	58.0	86	10	US-09-384-472-3	Sequence 3, Appli
	41	11.6	58.0	97	10	US-09-864-761-31748	Sequence 31748, A
c	42	11.4	57.0	54	10	US-09-051-034A-5	Sequence 5, Appli
	43	11.4	57.0	60	10	US-09-559-671A-46	Sequence 46, Appl
	44	11.2	56.0	18	9	US-09-978-295A-201	Sequence 201, App
	45	11.2	56.0	18	9	US-09-978-697-201	Sequence 201, App

ALIGNMENTS

RESULT 1
US-09-934-060A-8
; Sequence 8, Application US/09934060A
; Patent No. US20020155121A1
; GENERAL INFORMATION:
; APPLICANT: DeVico, Anthony L.
; APPLICANT: Fouts, Timothy R.
; APPLICANT: Tuskan, Robert G.
; TITLE OF INVENTION: VIRUS COAT PROTEIN/RECEPTOR CHIMERAS AND METHODS OF USE
; FILE REFERENCE: 4115-144 CIP
; CURRENT APPLICATION NUMBER: US/09/934,060A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/684,026
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/158,321
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthesized construct
US-09-934-060A-8

Query Match 71.0%; Score 14.2; DB 9; Length 66;
Best Local Similarity 84.2%; Pred. No. 2.5e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 GGGTCCGGAGCCTGAATCA 20
           ||||||||||| ||  ||
Db       1 GGGTCCGGAGCCCGAGCCA 19
```

RESULT 2
US-09-809-567-11/c
; Sequence 11, Application US/09809567
; Patent No. US20020045213A1
; GENERAL INFORMATION:
; APPLICANT: Jing, Shuqian
; TITLE OF INVENTION: IL-17 Receptor Like Molecules and Uses Thereof

```
;  FILE REFERENCE: 01017/36916A
;  CURRENT APPLICATION NUMBER: US/09/809,567
;  CURRENT FILING DATE:  2001-03-15
;  PRIOR APPLICATION NUMBER: 09/724,460
;  PRIOR FILING DATE: 2000-11-28
;  PRIOR APPLICATION NUMBER: 60/189,816
;  PRIOR FILING DATE: 2000-03-16
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 23
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer 2432-36
US-09-809-567-11

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 23;
  Best Local Similarity   83.3%;  Pred. No. 7.5e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATC 19
           | |||||||| ||| |||
Db      20 GAGTCCGGAGGCTGGATC 3


RESULT 3
US-09-864-761-20574/c
; Sequence 20574, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20574
;   LENGTH: 82
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AB016897.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.67
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.97
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.83
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.85
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
;   OTHER INFORMATION: NT HIT: AL161501.2, EVALUE 5.60e-01
;   OTHER INFORMATION: EST_HUMAN HIT: AW992151.1, EVALUE 2.60e+00
US-09-864-761-20574

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 82;
  Best Local Similarity   83.3%;  Pred. No. 7.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAAT 18
           |||||  | |||||||||
Db      63 TGGGTAGGCAGCCTGAAT 46


RESULT 4
US-09-886-055-519
; Sequence 519, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
;  APPLICANT: STRYER, LUBERT
;  APPLICANT:  ZOZULYA, SERGEY
;  TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
;  TITLE OF INVENTION:  BIOSENSORS OF CHEMICAL SENSANTS
;  FILE REFERENCE: 078003-0277150
;  CURRENT APPLICATION NUMBER: US/09/886,055
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 60/213,812
;  PRIOR FILING DATE: 2000-06-22
;  NUMBER OF SEQ ID NOS: 522
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 519
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Degenerate
;   OTHER INFORMATION: primer
;   NAME/KEY: modified_base
;   LOCATION: (20)
;   OTHER INFORMATION: A, T, C or G
;   NAME/KEY: modified_base
;   LOCATION: (22)..(23)
;   OTHER INFORMATION: A, T, C or G
US-09-886-055-519

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 1.2e+03;
  Matches   12;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAA 17
           |||||||||| :|:|:
Db       2 GGGTCCGGAGRSTRAD 17


RESULT 5
```

```
US-09-886-055-521
; Sequence 521, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
;  APPLICANT: STRYER, LUBERT
;  APPLICANT:  ZOZULYA, SERGEY
;  TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
;  TITLE OF INVENTION:  BIOSENSORS OF CHEMICAL SENSANTS
;  FILE REFERENCE: 078003-0277150
;  CURRENT APPLICATION NUMBER: US/09/886,055
;  CURRENT FILING DATE:  2001-06-22
;  PRIOR APPLICATION NUMBER: 60/213,812
;  PRIOR FILING DATE: 2000-06-22
;  NUMBER OF SEQ ID NOS: 522
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 521
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Degenerate
;   OTHER INFORMATION: primer
;   NAME/KEY: modified_base
;   LOCATION: (20)
;   OTHER INFORMATION: A, T, C or G
;   NAME/KEY: modified_base
;   LOCATION: (22)..(23)
;   OTHER INFORMATION: A, T, C or G
US-09-886-055-521

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 1.2e+03;
  Matches   12;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAA 17
           |||||||||| :|:|:
Db       2 GGGTCCGGAGRSTRAD 17


RESULT 6
US-09-780-668A-25/c
; Sequence 25, Application US/09780668A
; Patent No. US20020147311A1
; GENERAL INFORMATION:
;  APPLICANT: Gillies, Stephen
;  APPLICANT:  Burger, Christa
;  APPLICANT:  Lo, Kin-Ming
;  TITLE OF INVENTION: Enhancing the Circulating Half-Life of Antibody-Based Fusion
;  TITLE OF INVENTION:  Proteins
;  FILE REFERENCE: LEX-011
;  CURRENT APPLICATION NUMBER: US/09/780,668A
;  CURRENT FILING DATE:  2001-02-09
;  PRIOR APPLICATION NUMBER: US 60/181,768
;  PRIOR FILING DATE: 2000-02-11
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 5' oligonucleotide with linker
US-09-780-668A-25

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 33;
  Best Local Similarity   78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATCA 20
           || |||||| |||||  ||
Db      31 GGATCCGGACCCTGATCCA 13


RESULT 7
US-09-780-668A-26
; Sequence 26, Application US/09780668A
; Patent No. US20020147311A1
; GENERAL INFORMATION:
;  APPLICANT: Gillies, Stephen
;  APPLICANT:  Burger, Christa
;  APPLICANT:  Lo, Kin-Ming
;  TITLE OF INVENTION: Enhancing the Circulating Half-Life of Antibody-Based Fusion
;  TITLE OF INVENTION:  Proteins
;  FILE REFERENCE: LEX-011
;  CURRENT APPLICATION NUMBER: US/09/780,668A
;  CURRENT FILING DATE:  2001-02-09
;  PRIOR APPLICATION NUMBER: US 60/181,768
;  PRIOR FILING DATE: 2000-02-11
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 3' oligonucleotide with linker
US-09-780-668A-26

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 33;
  Best Local Similarity   78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATCA 20
           || |||||| |||||  ||
Db       3 GGATCCGGACCCTGATCCA 21


RESULT 8
US-09-821-821-17/c
; Sequence 17, Application US/09821821
; Patent No. US20020064823A1
; GENERAL INFORMATION:
;  APPLICANT: Welcher, Andrew A.
;  APPLICANT:  Calzone, Frank J.
;  TITLE OF INVENTION: CD20/IgE-Receptor Like Molecules and Uses Thereof
;  FILE REFERENCE: 01017/36938A
;  CURRENT APPLICATION NUMBER: US/09/821,821
;  CURRENT FILING DATE:  2001-03-29
;  PRIOR APPLICATION NUMBER: US 09/723,258
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: US 60/193,728
;  PRIOR FILING DATE: 2000-03-30
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 41
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer 2289-26
US-09-821-821-17

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 41;
  Best Local Similarity   78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAATC 19
           || || ||| || ||||||
Db      39 TGTGTGCGGTGCTTGAATC 21


RESULT 9
US-09-759-143-64
; Sequence 64, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
```

```
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 64
;   LENGTH: 97
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-759-143-64

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 97;
  Best Local Similarity   78.9%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAATC 19
          || ||| |||  |||||||
Db     12 TGAGTCGGGACGCTGAATC 30


RESULT 10
US-09-780-669-64
; Sequence 64, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 64
;   LENGTH: 97
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-780-669-64

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 97;
  Best Local Similarity   78.9%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAATC 19
          || ||| |||  |||||||
Db     12 TGAGTCGGGACGCTGAATC 30


RESULT 11
US-09-030-606-64
; Sequence 64, Application US/09030606
; Patent No. US20020081580A1
;  GENERAL INFORMATION:
;    APPLICANT:  Xu, Jiangchun
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS FOR IMMUNODIAGNOSIS OF PROSTATE CANCER AND METHOD
;    NUMBER OF SEQUENCES:  224
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/030,606
;      FILING DATE:  25-FEB-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.428C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  64:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  97 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-09-030-606-64

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 97;
  Best Local Similarity   78.9%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAATC 19
          || ||| |||  |||||||
Db     12 TGAGTCGGGACGCTGAATC 30


RESULT 12
US-09-822-827-64
; Sequence 64, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
```

```
;  FILE REFERENCE: 210121.534C1
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 64
;   LENGTH: 97
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-822-827-64

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 97;
  Best Local Similarity   78.9%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAATC 19
          || ||| |||  |||||||
Db     12 TGAGTCGGGACGCTGAATC 30


RESULT 13
US-09-115-453-64
; Sequence 64, Application US/09115453B
; Patent No. US20020090372A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT: Dillon, Davin C.
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
;  TITLE OF INVENTION: METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.427C4
;  CURRENT APPLICATION NUMBER: US/09/115,453B
;  CURRENT FILING DATE: 1998-07-14
;  NUMBER OF SEQ ID NOS: 228
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 64
;   LENGTH: 97
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-115-453-64

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 97;
  Best Local Similarity   78.9%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAATC 19
          || ||| |||  |||||||
Db     12 TGAGTCGGGACGCTGAATC 30


RESULT 14
US-10-081-281-16
; Sequence 16, Application US/10081281
; Patent No. US20020151707A1
;   GENERAL INFORMATION:
;        APPLICANT: Kindsvogel, Wayne
;                   Gross, Jane A.
;                   Sheppard, Paul
;        TITLE OF INVENTION: Immune Mediators and Related Methods
;        NUMBER OF SEQUENCES: 121
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,281
;             FILING DATE: 20-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/261,811A
;             FILING DATE: 03-Mar-1999
;             APPLICATION NUMBER: US 08/480,002
;             FILING DATE: 07-JUN-1995
;             APPLICATION NUMBER: US 08/482,133
;             FILING DATE: 07-JUN-1995
;             APPLICATION NUMBER: US 08/483,241
;             FILING DATE: 07-JUN-1995
;             APPLICATION NUMBER: US 60/005,964
;             FILING DATE: 27-OCT-1995
;             APPLICATION NUMBER: US 08/657,581
;             FILING DATE: 07-JUN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Parent, Annette S.
;             REGISTRATION NUMBER: 42,058
;             REFERENCE/DOCKET NUMBER: 014058-005630US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 576-0200
;             TELEFAX: (415) 576-0300
;   INFORMATION FOR SEQ ID NO: 16:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 39 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-081-281-16

  Query Match             62.0%;  Score 12.4;  DB 12;  Length 39;
  Best Local Similarity   92.9%;  Pred. No. 1.9e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCT 14
          ||||||||||| ||
Db     18 TGGGTCCGGAGACT 31


RESULT 15
US-09-969-708-357/c
; Sequence 357, Application US/09969708
; Patent No. US20020102532A1
; GENERAL INFORMATION:
;  APPLICANT: Augustus, Meena
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-70
;  CURRENT APPLICATION NUMBER: US/09/969,708
;  CURRENT FILING DATE:  2001-10-03
;  PRIOR APPLICATION NUMBER: US/60/237,606
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: US/60/237,608
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: US/60/237,425
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 658
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 357
;   LENGTH: 66
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-969-708-357

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 66;
  Best Local Similarity   92.9%;  Pred. No. 1.9e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GGGTCCGGAGCCTG 15
          |||||||||| |||
Db     61 GGGTCCGGAGGCTG 48
```

Search completed: November 23, 2002, 06:42:03
Job time : 18.25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-4
Perfect score: 20
Sequence: 1 tgggtccggagcctgaatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.2	71.0	88	12	BF117110	BF117110 uz05c06.y
c	2	13.8	69.0	30	12	BE735599	BE735599 601304856
c	3	13.8	69.0	55	9	AA498823	AA498823 vh32h02.r
c	4	13.8	69.0	58	9	AA869727	AA869727 vq15b03.r
c	5	13.8	69.0	60	9	AA080195	AA080195 mm33e01.r
c	6	13.8	69.0	82	14	W17669	W17669 mb76c07.r1
	7	13.8	69.0	90	9	AJ239904	AJ239904 AJ239904
	8	13.6	68.0	94	9	AA725453	AA725453 ai15h08.s
	9	13.4	67.0	77	14	N44726	N44726 yy37c09.r1
c	10	13.4	67.0	77	17	BH891526	BH891526 3526_1_18
c	11	13.2	66.0	65	9	AA096874	AA096874 mo02a03.r
c	12	13.2	66.0	68	13	BJ076607	BJ076607 BJ076607
c	13	13.2	66.0	75	9	AI902463	AI902463 CM-BT006-
c	14	12.8	64.0	45	13	BI519819	BI519819 603062250
	15	12.8	64.0	58	9	AA906833	AA906833 ok73f04.s
c	16	12.8	64.0	65	17	BH889500	BH889500 3526_1_10
	17	12.8	64.0	85	9	AA980524	AA980524 ua42c10.r
c	18	12.8	64.0	92	17	BH895490	BH895490 3526_1_34
	19	12.8	64.0	97	17	AG024646	AG024646 Oryza sat
c	20	12.8	64.0	100	14	BM820025	BM820025 K-EST0088
c	21	12.6	63.0	60	14	BM954016	BM954016 sam69a01.
c	22	12.6	63.0	78	9	AA836634	AA836634 od66b10.s
	23	12.6	63.0	81	10	AV742234	AV742234 AV742234
c	24	12.6	63.0	88	14	W16240	W16240 mb64b04.r1
	25	12.6	63.0	91	9	AA710526	AA710526 vt50f09.r
	26	12.6	63.0	92	17	BH852872	BH852872 SALK_0757
c	27	12.6	63.0	93	17	BH862040	BH862040 SALK_0885
	28	12.6	63.0	95	9	AA225611	AA225611 nc08e12.s
c	29	12.6	63.0	100	12	BG114826	BG114826 602315273
	30	12.4	62.0	59	9	AL798042	AL798042 AL798042
c	31	12.4	62.0	66	9	AA410714	AA410714 zv35h07.s
c	32	12.4	62.0	71	17	AZ816914	AZ816914 2M0085C21
	33	12.4	62.0	83	9	AA869319	AA869319 vq50e04.r
c	34	12.4	62.0	83	12	BG345720	BG345720 dg40d05.y
c	35	12.4	62.0	84	9	AA836964	AA836964 oe97g08.s
	36	12.4	62.0	85	13	BI906423	BI906423 603063464
	37	12.4	62.0	87	9	AA723245	AA723245 zg84c08.s
	38	12.4	62.0	91	9	AA621438	AA621438 af35g08.s
c	39	12.2	61.0	70	9	AI544943	AI544943 fb69f02.y
	40	12.2	61.0	75	14	H94092	H94092 yw58d11.r1
	41	12.2	61.0	76	9	AI793479	AI793479 fc48c01.x
	42	12.2	61.0	76	9	AA229247	AA229247 nc46g09.r
c	43	12.2	61.0	80	12	BF129926	BF129926 601817752
c	44	12.2	61.0	90	17	BH852734	BH852734 SALK_0755
c	45	12.2	61.0	94	14	H55585	H55585 CHR220524 C

ALIGNMENTS

RESULT 1
BF117110
LOCUS BF117110 88 bp mRNA linear EST 29-DEC-2000
DEFINITION uz05c06.y1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3668170 5' similar to SW:HV45_MOUSE P01821 IG HEAVY CHAIN V REGION MC101 PRECURSOR. ;, mRNA sequence.
ACCESSION BF117110
VERSION BF117110.1 GI:10986586
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 88)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MGI:1428938

```
          Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .88
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3668170"
                     /clone_lib="NCI_CGAP_Mam5"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
BASE COUNT       12 a     28 c     25 g     23 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 12;  Length 88;
  Best Local Similarity   84.2%;  Pred. No. 4.2e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGGTCCGGAGCCTGAATCA 20
          | ||||||||||  ||||
Db      3 GCGTCCGGAGCCTCCATCA 21


RESULT 2
BE735599/c
LOCUS       BE735599                 30 bp    mRNA    linear   EST 15-SEP-2000
DEFINITION  601304856F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3639421 5',
            mRNA sequence.
ACCESSION   BE735599
VERSION     BE735599.1  GI:10149591
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM341  row: p  column: 14.
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3639421"
                     /clone_lib="NIH_MGC_39"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Library constructed
                     by Ling Hong in the laboratory of Gerald M. Rubin
                     (University of California,  Berkeley) using ZAP-cDNA
                     synthesis kit (Stratagene) and Superscript II RT (Life
                     Technologies)."
BASE COUNT        5 a      7 c     11 g      7 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 12;  Length 30;
  Best Local Similarity   88.2%;  Pred. No. 5.8e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 GGTCCGGAGCCTGAATC 19
          || ||||| ||||||||
Db     17 GGGCCGGAACCTGAATC 1


RESULT 3
AA498823/c
LOCUS       AA498823                 55 bp    mRNA    linear   EST 01-JUL-1997
DEFINITION  vh32h02.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
            clone IMAGE:888723 5' similar to SW:RL35_HUMAN P42766 60S RIBOSOMAL
            PROTEIN L35. ;, mRNA sequence.
ACCESSION   AA498823
VERSION     AA498823.1  GI:2233846
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:516683
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:888723"
                     /clone_lib="Barstead mouse pooled organs MPLRB4"
                     /sex="mixed"
                     /tissue_type="pooled organs"
                     /dev_stage="7 day"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [GTTGGATTCGGTACC], digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead."
BASE COUNT       19 a     10 c     19 g      7 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 55;
  Best Local Similarity   88.2%;  Pred. No. 6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 GGTCCGGAGCCTGAATC 19
          ||||| |||||| ||||
Db     20 GGTCCCGAGCCTTAATC 4
```

```
RESULT 4
AA869727/c
LOCUS       AA869727                 58 bp    mRNA    linear   EST 16-MAR-1998
DEFINITION  vq15b03.r1 Barstead stromal cell line MPLRB8 Mus musculus cDNA
            clone IMAGE:1094285 5' similar to SW:RL35_RAT P17078 60S RIBOSOMAL
            PROTEIN L35. ;, mRNA sequence.
ACCESSION   AA869727
VERSION     AA869727.1  GI:2965172
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:600517
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1094285"
                     /clone_lib="Barstead stromal cell line MPLRB8"
                     /cell_line="C2C12 (undifferentiated)"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [AATTCGGATCCTTG], digested with Not I and cloned into the
                     Not I and Eco RI sites of the modified pT7T3 vector.
                     Source undifferentiated tissue culture cell line C2C12.
                     Library constructed by Bob Barstead. The C2C12 cell line
                     (available from ATCC, catalog # CRL-1772) differentiates
                     rapidly, forming contractile myotubes and producing
                     characteristic muscle proteins. "
BASE COUNT       19 a     10 c     21 g      8 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 9;  Length 58;
  Best Local Similarity  88.2%;  Pred. No. 6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GGTCCGGAGCCTGAATC 19
           ||||| |||||| ||||
Db      20 GGTCCCGAGCCTTAATC 4


RESULT 5
AA080195/c
LOCUS       AA080195                 60 bp    mRNA    linear   EST 06-FEB-1997
DEFINITION  mm33e01.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:523320 5', mRNA sequence.
ACCESSION   AA080195
VERSION     AA080195.1  GI:1619136
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:317168
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:523320"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       20 a     15 c     19 g      6 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 9;  Length 60;
  Best Local Similarity  88.2%;  Pred. No. 6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GGTCCGGAGCCTGAATC 19
           ||||| |||||| ||||
Db      44 GGTCCCGAGCCTTAATC 28


RESULT 6
W17669/c
LOCUS       W17669                   82 bp    mRNA    linear   EST 12-SEP-1996
DEFINITION  mb76c07.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
            IMAGE:335340 5' similar to SW:RL35_HUMAN P42766 60S RIBOSOMAL
            PROTEIN L35. ;, mRNA sequence.
ACCESSION   W17669
VERSION     W17669.1  GI:1538100
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
```

```
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1292054.
            Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:216740
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .82
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:335340"
                      /clone_lib="Soares mouse p3NMF19.5"
                      /dev_stage="19.5 dpc total fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Vector: pT7T3D (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. RNA was kindly provided by
                      Dr. Minoru Ko (Wayne State University)."
BASE COUNT       23 a     15 c     28 g     16 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 14;  Length 82;
  Best Local Similarity   88.2%;  Pred. No. 6.1e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GGTCCGGAGCCTGAATC 19
           ||||| |||||| ||||
Db      20 GGTCCCGAGCCTTAATC 4


RESULT 7
AJ239904
LOCUS       AJ239904                  90 bp    mRNA    linear   EST 10-AUG-1999
DEFINITION  AJ239904 Aspergillus niger ATCC6275 Aspergillus niger cDNA clone
            AN06A08, mRNA sequence.
ACCESSION   AJ239904
VERSION     AJ239904.1  GI:5443895
KEYWORDS    EST.
SOURCE      Aspergillus niger.
  ORGANISM  Aspergillus niger
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   Choi,J.Y., Lee,D.W., Koh,J.S., Kim,J.H., Yang,M.S. and Chae,K.S.
  TITLE     Identification of expressed sequence tags (ESTs) of the highly
            transcribed genes in Aspergillus niger
  JOURNAL   Biotechnol. Lett. 21, 381-384 (1999)
COMMENT     Contact: Chae KS
            Faculty of Biological Sciences
            Chonbuk National University
            Chonju 561-756, Republic of Korea.
FEATURES              Location/Qualifiers
     source           1. .90
                      /organism="Aspergillus niger"
                      /strain="ATCC6275"
                      /db_xref="taxon:5061"
                      /clone="AN06A08"
                      /clone_lib="Aspergillus niger ATCC6275"
BASE COUNT       26 a     19 c     25 g     19 t      1 others
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 90;
  Best Local Similarity   83.3%;  Pred. No. 6.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATC 19
           ||| ||||||| || |||
Db      26 GGGNCCGGAGCATGCATC 43


RESULT 8
AA725453
LOCUS       AA725453                  94 bp    mRNA    linear   EST 31-DEC-1998
DEFINITION  ai15h08.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
            1342911 3' similar to SW:NU1M_HUMAN P03886 NADH-UBIQUINONE
            OXIDOREDUCTASE CHAIN 1 ;, mRNA sequence.
ACCESSION   AA725453
VERSION     AA725453.1  GI:2743160
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
            , Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 644   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .94
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="1342911"
                      /clone_lib="Soares_parathyroid_tumor_NbHPA"
                      /tissue_type="parathyroid tumor"
                      /dev_stage="adult"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
                      ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                      RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                      primer
                      [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTT
                      TTTTT-3'], double-stranded cDNA was size selected, ligated
                      to Eco RI adapters (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of a modified pT7T3
                      vector (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
                      adenomas was kindly provided by Dr. Stephen Marx, National
                      Institute of Diabetes and Digestive and Kidney Diseases,
                      NIH."
```

```
BASE COUNT       24 a     25 c     23 g     22 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 94;
  Best Local Similarity   80.0%;  Pred. No. 7.5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGAATCA 20
           ||||||| | || | |||||
Db      11 TGGGTCCTGCGCGTAAATCA 30


RESULT 9
N44726
LOCUS       N44726                   77 bp    mRNA    linear   EST 13-FEB-1996
DEFINITION  yy37c09.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
            IMAGE:273424 5' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   N44726
VERSION     N44726.1  GI:1185892
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: T7
            High quality sequence stop: 65.
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Homo sapiens"
                     /db_xref="GDB:3883066"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:273424"
                     /clone_lib="Soares melanocyte 2NbHM"
                     /sex="Male"
                     /tissue_type="melanocyte"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. RNA from normal foreskin melanocytes
                     (FS374) was kindly provided by Dr. Anthony P. Albino."
BASE COUNT       16 a     31 c     22 g      6 t      2 others
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 14;  Length 77;
  Best Local Similarity   87.5%;  Pred. No. 8.9e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGGGTCCGGAGCCTGA 16
           |||| |||||||||  |
Db      18 TGGGCCCGGAGCCTNA 33


RESULT 10
BH891526/c
LOCUS       BH891526                 77 bp    DNA     linear   GSS 14-AUG-2002
DEFINITION  3526_1_18_1_B07.1EL_x_1 3526 - RescueMu Grid K Zea mays genomic,
            DNA sequence.
ACCESSION   BH891526
VERSION     BH891526.1  GI:22222561
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site of ends cut by single endonuclease.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 3526_1_18_1  row: 24
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="3526 - RescueMu Grid K"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid K was grown at Molokai, Hawaii in Winter
                     2000-2001. DNA was extracted from leaf punches, double
                     digested using BamHI and BglII, and ligated to form
                     circular plasmids. DH10B cells were transformed and then
                     screened on LB plates with ampicillin."
BASE COUNT       14 a     23 c     27 g     13 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 17;  Length 77;
  Best Local Similarity   93.3%;  Pred. No. 8.9e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 GGTCCGGAGCCTGAA 17
           || ||||||||||||
Db      42 GGCCCGGAGCCTGAA 28


RESULT 11
AA096874/c
LOCUS       AA096874                 65 bp    mRNA    linear   EST 15-FEB-1997
DEFINITION  mo02a03.r1 Stratagene mouse lung 937302 Mus musculus cDNA clone
            IMAGE:552364 5', mRNA sequence.
ACCESSION   AA096874
VERSION     AA096874.1  GI:1642585
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:333156
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 62.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Mus musculus"
                     /strain="C57BL/6 x CBA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:552364"
                     /clone_lib="Stratagene mouse lung 937302"
                     /sex="female"
                     /tissue_type="lung"
                     /dev_stage="6-8 month old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. 6-8 month old female lung and 1.5 year old male lung
                     were source of mRNA. Average insert size: 1.5 kb; Uni-ZAP
                     XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT        7 a     31 c     23 g      4 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 65;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGGTCCGGAGCCTGAATC 19
           ||| |||||| ||| |||
Db      59 GGGCCCGGAGGCTGCATC 42


RESULT 12
BJ076607/c
LOCUS       BJ076607                  68 bp    mRNA    linear   EST 11-DEC-2001
DEFINITION  BJ076607 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL058d03 3', mRNA sequence.
ACCESSION   BJ076607
VERSION     BJ076607.1  GI:17521523
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .68
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL058d03"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT       28 a     13 c     11 g     16 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 13;  Length 68;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GGTCCGGAGCCTGAATCA 20
           | | | ||||||||||||
Db      40 GTTTCAGAGCCTGAATCA 23


RESULT 13
AI902463/c
LOCUS       AI902463                  75 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  CM-BT006-170299-218 BT006 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI902463
VERSION     AI902463.1  GI:6492850
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=CM&t2=CM-BT006-218.html
            &t3=170299&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT006"
                     /sex="female"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
```

```
                ,716 -  Ludwig Institute for Cancer Research) profiles
                into  the pUC 18 vector. Reverse transcription of tissue
                mRNA and cDNA amplification were performed under low
                stringency conditions."
BASE COUNT       17 a     19 c     21 g     12 t      6 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 75;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGGGTCCGGAGCCTGAAT 18
          ||||||| ||||  ||||
Db     72 TGGGTCCCGAGCTCGAAT 55


RESULT 14
BI519819/c
LOCUS       BI519819                 45 bp    mRNA    linear   EST 29-AUG-2001
DEFINITION  603062250T1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5211228 3',
            mRNA sequence.
ACCESSION   BI519819
VERSION     BI519819.1  GI:15344611
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11530  row: d  column: 13
            High quality sequence stop: 45.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5211228"
                     /clone_lib="NIH_MGC_118"
                     /tissue_type="leukocyte"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                     (destroyed); RNA source leukocytes from anonymous pool of
                     non-activated  adult donors.  Library is oligo-dT primed
                     and directionally cloned  (EcoRV site is destroyed upon
                     cloning).  Average insert size 1.7 kb, insert size range
                     1.2-3.3 kb.  Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking code 027. Note:
                     this is a NIH_MGC Library."
BASE COUNT        8 a     11 c     17 g      9 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 13;  Length 45;
  Best Local Similarity   87.5%;  Pred. No. 1.5e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 TCCGGAGCCTGAATCA 20
          ||| ||||||| ||||
Db     28 TCCAGAGCCTGGATCA 13


RESULT 15
AA906833
LOCUS       AA906833                 58 bp    mRNA    linear   EST 09-JUN-1998
DEFINITION  ok73f04.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1519615 3'
            similar to SW:NHPX_HUMAN P55769 NHP2/RS6 FAMILY PROTEIN YEL026W
            HOMOLOG ;, mRNA sequence.
ACCESSION   AA906833
VERSION     AA906833.1  GI:3042077
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 599   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1519615"
                     /clone_lib="NCI_CGAP_GC4"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from 3 pooled
                     germ cell tumors, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       10 a     13 c     11 g     24 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 9;  Length 58;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 TCCGGAGCCTGAATCA 20
          |||||||||||| || |
Db     37 TCCGGAGCCTGTATAA 52


Search completed: November 26, 2002, 04:06:50
Job time : 767.8 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-5
Perfect score: 20
Sequence: 1 tttttcagggaatccggggg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	14.2	71.0	52	5	TCUMITO3	AF154453 Toxostoma
2	14.2	71.0	75	9	AB00468S46	AB004622 Human gen
c 3	13.8	69.0	53	5	TCIMITO3	AF154459 Toxostoma
4	13.6	68.0	50	6	AX160046	AX160046 Sequence
5	13.6	68.0	51	6	AX160045	AX160045 Sequence
c 6	13.6	68.0	54	5	TO1MITO3	AF154447 Toxostoma
c 7	13.6	68.0	97	11	AL772837	AL772837 Arabidops
8	13.4	67.0	21	6	AR020977	AR020977 Sequence
c 9	13.4	67.0	21	6	AR020978	AR020978 Sequence
10	13.4	67.0	21	6	AR043392	AR043392 Sequence
c 11	13.4	67.0	21	6	AR043393	AR043393 Sequence
12	13.4	67.0	21	6	AR062307	AR062307 Sequence
c 13	13.4	67.0	21	6	AR062308	AR062308 Sequence
14	13.4	67.0	21	6	AR183766	AR183766 Sequence
c 15	13.4	67.0	21	6	AR183767	AR183767 Sequence
16	13.2	66.0	36	6	AR061022	AR061022 Sequence
17	13.2	66.0	36	6	AX018732	AX018732 Sequence
c 18	13.2	66.0	53	5	TREMITO3	AF154483 Toxostoma
c 19	13.2	66.0	53	5	TLE3MITO3	AF154489 Toxostoma
c 20	13.2	66.0	53	5	TLE2MITO3	AF154495 Toxostoma
21	13.2	66.0	59	9	HUMACXIXC2	D38461 Homo sapien
c 22	12.8	64.0	22	6	AX365220	AX365220 Sequence
c 23	12.8	64.0	28	6	AX229851	AX229851 Sequence
24	12.8	64.0	51	6	AX162296	AX162296 Sequence
25	12.8	64.0	66	9	AF189424	AF189424 Homo sapi
c 26	12.8	64.0	91	1	AF087267	AF087267 Chlamydia
c 27	12.6	63.0	24	6	AX445637	AX445637 Sequence
28	12.6	63.0	28	6	AX252923	AX252923 Sequence
29	12.6	63.0	28	6	AX342689	AX342689 Sequence
30	12.6	63.0	28	6	AX374942	AX374942 Sequence
31	12.6	63.0	80	9	HUMIGHZCJ	L23292 Human rearr
32	12.6	63.0	84	9	AB00468S30	AB004605 Human gen
33	12.6	63.0	88	14	PVXANPTII	X62443 Potato viru
34	12.4	62.0	76	3	TA28SRNAB	Z18252 Tritrichomo
35	12.4	62.0	78	3	HA28SRNAB	Z18248 Hypotrichom
36	12.4	62.0	78	3	TV28SRNAC	Z18241 Trichomonas
37	12.4	62.0	79	3	PH28SRNAA	Z18250 Pentatricho
c 38	12.4	62.0	81	6	AR156804	AR156804 Sequence
c 39	12.4	62.0	85	10	MUSPKCD02	M18241 Mouse cAMP-
c 40	12.2	61.0	19	6	A38742	A38742 Sequence 53
c 41	12.2	61.0	19	6	AR102156	AR102156 Sequence
42	12.2	61.0	19	6	AX131000	AX131000 Sequence
c 43	12.2	61.0	20	4	DOGP45301	L24338 Dog (Clone:
44	12.2	61.0	20	6	AX023313	AX023313 Sequence
45	12.2	61.0	21	6	AX058986	AX058986 Sequence

ALIGNMENTS

RESULT 1
TCUMITO3/c
LOCUS TCUMITO3 52 bp DNA linear VRT 23-SEP-1999
DEFINITION Toxostoma curvirostre isolate CBTH6TX tRNA-Glu gene, partial sequence; mitochondrial gene for mitochondrial product.
ACCESSION AF154453
VERSION AF154453.1 GI:5918791
KEYWORDS .
SEGMENT 3 of 6
SOURCE Toxostoma curvirostre.
ORGANISM Mitochondrion Toxostoma curvirostre
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Passeriformes; Sturnidae; Toxostoma.
REFERENCE 1 (bases 1 to 52)
AUTHORS Zink,R.M., Dittmann,D.L., Klicka,J. and Blackwell-Rago,R.C.

```
  TITLE     Evolutionary patterns of morphometrics, allozymes and mitochondrial
            DNA in thrashers (Toxostoma)
  JOURNAL   The Auk (1999) In press
REFERENCE   2  (bases 1 to 52)
  AUTHORS   Zink,R.M., Dittmann,D.L., Klicka,J. and Blackwell-Rago,R.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAY-1999) Bell Museum, University of Minnesota, 100
            Ecology Bldg, 1987 Upper Buford Circle, St Paul, MN 55108, USA
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Toxostoma curvirostre"
                     /organelle="mitochondrion"
                     /isolate="CBTH6TX"
                     /specimen_voucher="RMZ 2951, LSU Museum of Zoology"
                     /db_xref="taxon:99878"
                     /sex="male"
                     /country="USA:Texas, 10 miles SSE of Charlotte"
                     /note="isolated May 26, 1992"
     tRNA            <1. .>52
                     /product="tRNA-Glu"
BASE COUNT       12 a     12 c      9 g     19 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 5;  Length 52;
  Best Local Similarity   84.2%;  Pred. No. 1.6e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCGGGGG 20
          ||||||| ||||| || |
Db     32 TTTTCAGGAAATCCTGGAG 14


RESULT 2
AB00468S46
LOCUS       AB00468S46               75 bp    DNA     linear   PRI 14-APR-2000
DEFINITION  Human gene for Type XIX collagen a1 chain, exon 46.
ACCESSION   AB004622
VERSION     AB004622.1  GI:2598711
KEYWORDS    Type XIX collagen a1 chain.
SEGMENT     46 of 52
SOURCE      Homo sapiens DNA, clone:KIG12.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Khaleduzzaman,M., Sumiyoshi,H., Ueki,Y., Inoguchi,K., Ninomiya,Y.
            and Yoshioka,H.
  TITLE     Structure of the human type XIX collagen (COL19A1) gene, which
            suggests it has arisen from an ancestor gene of the FACIT family
  JOURNAL   Genomics 45 (2), 304-312 (1997)
  MEDLINE   98008918
REFERENCE   2  (bases 1 to 75)
  AUTHORS   Yoshioka,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-1997) Hidekatsu Yoshioka, Okayama University
            Medical School, Molecular Biology and Biochemistry; 2-5-1
            Shikata-cho, Okayama, Okayama 700, Japan
            (E-mail:mbb@ccews.cc.okayama-u.ac.jp, Tel:086-235-7128,
            Fax:o86-222-7768)
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="KIG12"
     exon            16. .60
                     /gene="COL19A1"
                     /number=46
BASE COUNT       22 a     14 c     20 g     19 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 75;
  Best Local Similarity   84.2%;  Pred. No. 1.6e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGG 19
          |||||||||| |||  |||
Db      8 TTTTTCAGGGCATCATGGG 26


RESULT 3
TCIMITO3/c
LOCUS       TCIMITO3                 53 bp    DNA     linear   VRT 23-SEP-1999
DEFINITION  Toxostoma cinereum isolate GRTH4BN tRNA-Glu gene, partial sequence;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF154459
VERSION     AF154459.1  GI:5918797
KEYWORDS    .
SEGMENT     3 of 6
SOURCE      Toxostoma cinereum.
  ORGANISM  Mitochondrion Toxostoma cinereum
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Passeriformes; Sturnidae; Toxostoma.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Zink,R.M., Dittmann,D.L., Klicka,J. and Blackwell-Rago,R.C.
  TITLE     Evolutionary patterns of morphometrics, allozymes and mitochondrial
            DNA in thrashers (Toxostoma)
  JOURNAL   The Auk (1999) In press
REFERENCE   2  (bases 1 to 53)
  AUTHORS   Zink,R.M., Dittmann,D.L., Klicka,J. and Blackwell-Rago,R.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAY-1999) Bell Museum, University of Minnesota, 100
            Ecology Bldg, 1987 Upper Buford Circle, St Paul, MN 55108, USA
FEATURES             Location/Qualifiers
     source          1. .53
                     /organism="Toxostoma cinereum"
                     /organelle="mitochondrion"
                     /isolate="GRTH4BN"
                     /specimen_voucher="B16587, LSU Museum of Zoology"
                     /specimen_voucher="RMZ 2165, American Museum, New York"
                     /db_xref="taxon:99877"
                     /sex="female"
                     /country="Mexico:Baja California, 3 km. west of Camalu"
                     /note="isolated November 10, 1983"
     tRNA            <1. .>53
                     /product="tRNA-Glu"
BASE COUNT       13 a     14 c      9 g     17 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 5;  Length 53;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCGGG 18
          |||||||| ||||| ||
Db     33 TTTTCAGGAAATCCTGG 17


RESULT 4
AX160046
LOCUS       AX160046                 50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3374 from Patent WO0140521.
ACCESSION   AX160046
VERSION     AX160046.1  GI:14541377
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 3374 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg43253001"
     misc_feature    26
                     /note="2 of 2 allelic variants (3373 is other entry)"
BASE COUNT       10 a     12 c     19 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          | | || ||||| |||||||
Db      4 TGTCTCGGGGAAGCCGGGGG 23


RESULT 5
AX160045
LOCUS       AX160045                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3373 from Patent WO0140521.
ACCESSION   AX160045
VERSION     AX160045.1  GI:14541376
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 3373 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (3374 is other entry)
                     Accession number cg43253001"
BASE COUNT       10 a     12 c     20 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          | | || ||||| |||||||
Db      4 TGTCTCGGGGAAGCCGGGGG 23


RESULT 6
TO1MITO3/c
LOCUS       TO1MITO3                  54 bp    DNA     linear   VRT 23-SEP-1999
DEFINITION  Toxostoma ocellatum isolate OCTH2MX tRNA-Glu gene, partial
            sequence; mitochondrial gene for mitochondrial product.
ACCESSION   AF154447
VERSION     AF154447.1  GI:5918785
KEYWORDS    .
SEGMENT     3 of 6
SOURCE      Toxostoma ocellatum.
  ORGANISM  Mitochondrion Toxostoma ocellatum
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Passeriformes; Sturnidae; Toxostoma.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Zink,R.M., Dittmann,D.L., Klicka,J. and Blackwell-Rago,R.C.
  TITLE     Evolutionary patterns of morphometrics, allozymes and mitochondrial
            DNA in thrashers (Toxostoma)
  JOURNAL   The Auk (1999) In press
REFERENCE   2  (bases 1 to 54)
  AUTHORS   Zink,R.M., Dittmann,D.L., Klicka,J. and Blackwell-Rago,R.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAY-1999) Bell Museum, University of Minnesota, 100
            Ecology Bldg, 1987 Upper Buford Circle, St Paul, MN 55108, USA
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Toxostoma ocellatum"
                     /organelle="mitochondrion"
                     /isolate="OCTH2MX"
                     /specimen_voucher="RMZ 2190, American Museum, New York"
                     /specimen_voucher="B16570, LSU Museum of Zoology"
                     /db_xref="taxon:99881"
                     /sex="male"
                     /country="Mexico:Oaxaca, 5 km SE of Matatlan"
                     /note="isolated November 18, 1993"
     tRNA            <1. .>54
                     /product="tRNA-Glu"
BASE COUNT       15 a     15 c      7 g     17 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 5;  Length 54;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          |||||||||  |||| || |
Db     34 TTTTTCAGGAGATCCTGGAG 15


RESULT 7
AL772837/c
LOCUS       AL772837                  97 bp    DNA     linear   STS 19-JUN-2002
DEFINITION  Arabidopsis thaliana transposon insertion STS SM_3.805, sequence
            tagged site.
ACCESSION   AL772837
VERSION     AL772837.1  GI:21535024
KEYWORDS    STS; STS, sequence tagged site.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B.,
            Murphy,G., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 97)
  AUTHORS   Clarke,J.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUN-2002) Clarke J.H., John Innes Centre, Colney
            Lane, Norwich, NR4 7UJ, UK
COMMENT     AT denotes an activation tag dissociation transposon within a
            single line, ET an enhancer trap dissociation transposon, GT a gene
            trap dissociation transposon, MT a mis-expression enhancer trap
            dissociation transposon, SM a defective suppressor mutator
            transposon. _3 denotes a sequence derived from the 3'end of the
            transposon, _5 denotes a sequence derived from the 5'end of the
            transposon BBSRC GARNet, ATIS project
            On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
            code: N55688.
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="Arabidopsis thaliana"
                     /variety="Landsberg erecta"
                     /db_xref="taxon:3702"
                     /clone="AC006223"
     STS             1. .97
                     /standard_name="SM_3.805"
BASE COUNT       43 a     17 c     13 g     24 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 11;  Length 97;
```

```
  Best Local Similarity   80.0%;  Pred. No. 3.2e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          |||||  ||  |||||||||
Db     87 TTTTTTTGGCCATCCGGGGG 68


RESULT 8
AR020977
LOCUS       AR020977                21 bp    DNA     linear   PAT 05-DEC-1998
DEFINITION  Sequence 9 from patent US 5789245.
ACCESSION   AR020977
VERSION     AR020977.1  GI:3975592
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dubensky,T.W. Jr., Polo,J.M., Ibanez,C.E., Chang,S.M.W., Jolly,D.J.
            and Driver,D.A.
  TITLE     Alphavirus structural protein expression cassettes
  JOURNAL   Patent: US 5789245-A 9 04-AUG-1998;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        3 a      8 c      5 g      5 t
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 6;  Length 21;
  Best Local Similarity  93.3%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCG 16
          |||||||||  |||||
Db      3 TTTTCAGGGGATCCG 17


RESULT 9
AR020978/c
LOCUS       AR020978                21 bp    DNA     linear   PAT 05-DEC-1998
DEFINITION  Sequence 10 from patent US 5789245.
ACCESSION   AR020978
VERSION     AR020978.1  GI:3975593
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dubensky,T.W. Jr., Polo,J.M., Ibanez,C.E., Chang,S.M.W., Jolly,D.J.
            and Driver,D.A.
  TITLE     Alphavirus structural protein expression cassettes
  JOURNAL   Patent: US 5789245-A 10 04-AUG-1998;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        5 a      5 c      8 g      3 t
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 6;  Length 21;
  Best Local Similarity  93.3%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCG 16
          |||||||||  |||||
Db     19 TTTTCAGGGGATCCG 5


RESULT 10
AR043392
LOCUS       AR043392                21 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 9 from patent US 5814482.
ACCESSION   AR043392
VERSION     AR043392.1  GI:5964400
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dubensky,T.W. Jr., Polo,J.M., Jolly,D.J. and Driver,D.A.
  TITLE     Eukaryotic layered vector initiation systems
  JOURNAL   Patent: US 5814482-A 9 29-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        3 a      8 c      5 g      5 t
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 6;  Length 21;
  Best Local Similarity  93.3%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCG 16
          |||||||||  |||||
Db      3 TTTTCAGGGGATCCG 17


RESULT 11
AR043393/c
LOCUS       AR043393                21 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 10 from patent US 5814482.
ACCESSION   AR043393
VERSION     AR043393.1  GI:5964401
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dubensky,T.W. Jr., Polo,J.M., Jolly,D.J. and Driver,D.A.
  TITLE     Eukaryotic layered vector initiation systems
  JOURNAL   Patent: US 5814482-A 10 29-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        5 a      5 c      8 g      3 t
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 6;  Length 21;
  Best Local Similarity  93.3%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCG 16
          |||||||||  |||||
Db     19 TTTTCAGGGGATCCG 5


RESULT 12
AR062307
LOCUS       AR062307                21 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 9 from patent US 5843723.
ACCESSION   AR062307
VERSION     AR062307.1  GI:5989998
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dubensky,T.W. Jr., Polo,J.M., Ibanez,C.E., Chang,S.M.W.,
            Jolly,D.J., Driver,D.A. and Belli,B.A.
  TITLE     Alphavirus vector constructs
  JOURNAL   Patent: US 5843723-A 9 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        3 a      8 c      5 g      5 t
ORIGIN
```

```
  Query Match             67.0%;  Score 13.4;  DB 6;  Length 21;
  Best Local Similarity   93.3%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCG 16
          ||||||||| |||||
Db      3 TTTTCAGGGGATCCG 17


RESULT 13
AR062308/c
LOCUS       AR062308                  21 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 10 from patent US 5843723.
ACCESSION   AR062308
VERSION     AR062308.1  GI:5989999
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dubensky,T.W. Jr., Polo,J.M., Ibanez,C.E., Chang,S.M.W.,
            Jolly,D.J., Driver,D.A. and Belli,B.A.
  TITLE     Alphavirus vector constructs
  JOURNAL   Patent: US 5843723-A 10 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        5 a      5 c      8 g      3 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 21;
  Best Local Similarity   93.3%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCG 16
          ||||||||| |||||
Db     19 TTTTCAGGGGATCCG 5


RESULT 14
AR183766
LOCUS       AR183766                  21 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 9 from patent US 6342372.
ACCESSION   AR183766
VERSION     AR183766.1  GI:20227735
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dubensky,T.W. Jr., Polo,J.M. and Driver,D.A.
  TITLE     Eukaryotic layered vector initiation systems for production of
            recombinant proteins
  JOURNAL   Patent: US 6342372-A 9 29-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        3 a      8 c      5 g      5 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 21;
  Best Local Similarity   93.3%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCG 16
          ||||||||| |||||
Db      3 TTTTCAGGGGATCCG 17


RESULT 15
AR183767/c
LOCUS       AR183767                  21 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 10 from patent US 6342372.
ACCESSION   AR183767
VERSION     AR183767.1  GI:20227736
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dubensky,T.W. Jr., Polo,J.M. and Driver,D.A.
  TITLE     Eukaryotic layered vector initiation systems for production of
            recombinant proteins
  JOURNAL   Patent: US 6342372-A 10 29-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        5 a      5 c      8 g      3 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 21;
  Best Local Similarity   93.3%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCG 16
          ||||||||| |||||
Db     19 TTTTCAGGGGATCCG 5


Search completed: December  3, 2002, 18:12:27
Job time : 311.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-5
Perfect score: 20
Sequence: 1 tttttcagggaatccggggg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		12.6	63.0	23	10	US-09-303-510-40	Sequence 40, Appl
2		12.6	63.0	23	10	US-09-303-040-40	Sequence 40, Appl
3		12.6	63.0	69	10	US-09-785-632A-44	Sequence 44, Appl
	c 4	12.2	61.0	36	10	US-09-749-234A-27	Sequence 27, Appl
	c 5	12.2	61.0	87	10	US-09-923-876-6135	Sequence 6135, Ap
	c 6	12.2	61.0	88	10	US-09-783-590-4956	Sequence 4956, Ap
7		12.2	61.0	88	10	US-09-783-590-7099	Sequence 7099, Ap
	c 8	12.2	61.0	89	10	US-09-878-574-6486	Sequence 6486, Ap
	c 9	12.2	61.0	100	10	US-09-864-761-24917	Sequence 24917, A
	c 10	12	60.0	40	10	US-09-759-143-850	Sequence 850, App
	c 11	12	60.0	40	10	US-09-780-669-850	Sequence 850, App
	c 12	12	60.0	40	10	US-09-822-827-850	Sequence 850, App
13		12	60.0	84	10	US-09-864-761-32573	Sequence 32573, A
14		12	60.0	86	10	US-09-864-761-32016	Sequence 32016, A
	c 15	11.8	59.0	26	10	US-09-736-169-2	Sequence 2, Appli
	c 16	11.8	59.0	26	12	US-10-055-174-2	Sequence 2, Appli
	c 17	11.8	59.0	69	10	US-09-764-847-367	Sequence 367, App
18		11.8	59.0	97	10	US-09-294-093B-2105	Sequence 2105, Ap
	c 19	11.6	58.0	31	10	US-09-870-122-8	Sequence 8, Appli
	c 20	11.6	58.0	36	10	US-09-006-298-2	Sequence 2, Appli
21		11.6	58.0	45	10	US-09-781-902-59	Sequence 59, Appl
	c 22	11.6	58.0	48	9	US-09-900-425A-29	Sequence 29, Appl
23		11.6	58.0	50	10	US-09-892-864A-26	Sequence 26, Appl
24		11.6	58.0	50	10	US-09-996-561-27	Sequence 27, Appl
	c 25	11.6	58.0	68	10	US-09-783-590-1427	Sequence 1427, Ap
26		11.6	58.0	74	10	US-09-974-300-4065	Sequence 4065, Ap
27		11.6	58.0	86	10	US-09-864-761-31082	Sequence 31082, A
	c 28	11.4	57.0	18	10	US-09-789-556A-41	Sequence 41, Appl
	c 29	11.4	57.0	31	10	US-09-801-274-141	Sequence 141, App
	c 30	11.4	57.0	32	10	US-09-769-864-50	Sequence 50, Appl
	c 31	11.4	57.0	38	10	US-09-972-331-7	Sequence 7, Appli
	c 32	11.4	57.0	48	10	US-09-766-378A-15	Sequence 15, Appl
	c 33	11.4	57.0	52	10	US-09-978-274A-13	Sequence 13, Appl
34		11.4	57.0	55	10	US-09-965-099-41	Sequence 41, Appl
35		11.4	57.0	55	12	US-10-051-852-41	Sequence 41, Appl
36		11.4	57.0	93	10	US-09-284-663A-28	Sequence 28, Appl
	c 37	11.2	56.0	25	10	US-09-784-911-18	Sequence 18, Appl
38		11.2	56.0	27	9	US-09-731-449-60	Sequence 60, Appl
39		11.2	56.0	31	10	US-09-801-274-505	Sequence 505, App
	c 40	11.2	56.0	39	10	US-09-859-214-63	Sequence 63, Appl
	c 41	11.2	56.0	42	9	US-09-924-647-5	Sequence 5, Appli
	c 42	11.2	56.0	42	9	US-09-905-291A-61	Sequence 61, Appl
	c 43	11.2	56.0	42	10	US-09-767-609-5	Sequence 5, Appli
	c 44	11.2	56.0	42	10	US-09-909-320-61	Sequence 61, Appl
	c 45	11.2	56.0	42	10	US-09-909-088B-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1
US-09-303-510-40
; Sequence 40, Application US/09303510A
; Patent No. US20020028208A1
; GENERAL INFORMATION:
; APPLICANT: Collisson, Ellen W.
; APPLICANT: Hash, Stephen M.
; APPLICANT: Choi, InSoo
; TITLE OF INVENTION: Feline CD80, Feline CD86, Feline CD28, and Feline
; TITLE OF INVENTION: CTLA-4 Nucleic Acid and Polypeptides
; FILE REFERENCE: 54954
; CURRENT APPLICATION NUMBER: US/09/303,510A
; CURRENT FILING DATE: 1999-04-30
; EARLIER APPLICATION NUMBER: 60/083,869
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer
US-09-303-510-40

Query Match 63.0%; Score 12.6; DB 10; Length 23;
Best Local Similarity 78.9%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 TTTTTCAGGGAATCCGGGG 19
           |||||||||  | | ||||
Db       5 TTTTTCAGGTTAGCAGGGG 23
```

RESULT 2
US-09-303-040-40
; Sequence 40, Application US/09303040
; Patent No. US20020051792A1
; GENERAL INFORMATION:
; APPLICANT: Winslow, Barbara J.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: Recombinant Virus Expressing Foreign DNA Encoding

```
;  TITLE OF INVENTION: Feline CD80, Feline CD86, Feline CD28, Feline CTLA-4 or
;  TITLE OF INVENTION: Feline Interferon-gama And Uses Thereof
;  FILE REFERENCE: 54957-B
;  CURRENT APPLICATION NUMBER: US/09/303,040
;  CURRENT FILING DATE: 1999-04-30
;  EARLIER APPLICATION NUMBER: 60/083,870
;  EARLIER FILING DATE: 1998-05-01
;  NUMBER OF SEQ ID NOS: 82
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
;    LENGTH: 23
;    TYPE: DNA
;    ORGANISM: feline CD86 primer
US-09-303-040-40

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 23;
  Best Local Similarity   78.9%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TTTTTCAGGGAATCCGGGG 19
           |||||||||  | | ||||
Db       5 TTTTTCAGGTTAGCAGGGG 23


RESULT 3
US-09-785-632A-44
; Sequence 44, Application US/09785632A
; Patent No. US20020061512A1
; GENERAL INFORMATION:
;  APPLICANT: Kim, Jin-Soo
;  APPLICANT:  Kwon, Young Do
;  APPLICANT:  Kim, Hyun-Won
;  APPLICANT:  Ryu, Eun-Hyun
;  APPLICANT:  Hwang, Moon-Sun
;  TITLE OF INVENTION: ZINC FINGER DOMAINS AND METHODS OF
;  TITLE OF INVENTION:  IDENTIFYING SAME
;  FILE REFERENCE: 12279-002001
;  CURRENT APPLICATION NUMBER: US/09/785,632A
;  CURRENT FILING DATE:  2001-02-16
;  NUMBER OF SEQ ID NOS: 166
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
;    LENGTH: 69
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (1)...(69)
US-09-785-632A-44

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 69;
  Best Local Similarity   78.9%;  Pred. No. 1.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TTTTCAGGGAATCCGGGGG 20
           ||| ||||||||  ||| |
Db       5 TTTGCAGGGAATGTGGGCG 23


RESULT 4
US-09-749-234A-27/c
; Sequence 27, Application US/09749234A
; Patent No. US20020068817A1
;   GENERAL INFORMATION:
;        APPLICANT: HOFMANN, Joachim
;                   SCHMID, Karlheinrich
;                   PAULI, Annette
;        TITLE OF INVENTION: DICTYOCAULUS VIVIPARUS ANTIGEN FOR THE
;                            DIAGNOSIS OF LUNGWORM INFESTATION AND FOR VACCINATION
;        NUMBER OF SEQUENCES: 30
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: FOLEY & LARDNER
;             STREET: 3000 K Street, N.W.
;             CITY: Washington
;             STATE: D.C.
;             COUNTRY: U.S.A.
;             ZIP: 20007-5109
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/749,234A
;             FILING DATE: 27-Dec-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/403,092
;             FILING DATE: 199-10-15
;             APPLICATION NUMBER: DE 197 15 586.3
;             FILING DATE: 15-APR-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Granados, Patricia D.
;             REGISTRATION NUMBER: 33,683
;             REFERENCE/DOCKET NUMBER: 038311/0103
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (202) 672-5300
;             TELEFAX: (202) 672-5399
;   INFORMATION FOR SEQ ID NO: 27:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 36 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        FEATURE:
;             NAME/KEY:  exon
;             LOCATION:  1..36
;        SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-749-234A-27

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 36;
  Best Local Similarity   82.4%;  Pred. No. 1.9e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TTTTCAGGGAATCCGGG 18
           | ||||||||||| | ||
Db      17 TGTTCAGGGAATTCCGG 1


RESULT 5
US-09-923-876-6135/c
; Sequence 6135, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
;  APPLICANT: Lalgudi, Raghunath V.
;  APPLICANT:  Kamigaki, Laura Y. (Ito)
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
;  FILE REFERENCE: PL-0012-1 CON
;  CURRENT APPLICATION NUMBER: US/09/923,876
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: 09/298,329
;  PRIOR FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: 60/085,331
;  PRIOR FILING DATE: 1998-05-05
;  NUMBER OF SEQ ID NOS: 6332
;  SOFTWARE: PERL Program
; SEQ ID NO 6135
;    LENGTH: 87
;    TYPE: DNA
;    ORGANISM: Zea mays
;    FEATURE:
;    NAME/KEY: misc_feature
;    OTHER INFORMATION: Incyte ID No. US20020013958A1 700458438H1
;    NAME/KEY: unsure
```

```
;   LOCATION: 31, 47, 59, 86
;   OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-6135

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 87;
  Best Local Similarity   77.8%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      3 TTTCAGGGAATCCGGGGG 20
          ||||   |||||||||  |
Db     52 TTTCCNCGAATCCGGGTG 35


RESULT 6
US-09-783-590-4956/c
; Sequence 4956, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
;  APPLICANT: Dillon, Patrick J.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Li, Haodong
;  APPLICANT:  Rosen, Craig A.
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
;  FILE REFERENCE: PO-16.2C1
;  CURRENT APPLICATION NUMBER: US/09/783,590
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 08/420,856
;  PRIOR FILING DATE: 1995-04-12
;  PRIOR APPLICATION NUMBER: 08/346,731
;  PRIOR FILING DATE: 1994-11-21
;  NUMBER OF SEQ ID NOS: 12485
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4956
;   LENGTH: 88
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (15)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (86)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-4956

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 88;
  Best Local Similarity   77.8%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGG 18
          |||||  |||  |  |  ||||
Db     20 TTTTTNAGGAACTGCGGG 3


RESULT 7
US-09-783-590-7099
; Sequence 7099, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
;  APPLICANT: Dillon, Patrick J.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Li, Haodong
;  APPLICANT:  Rosen, Craig A.
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
;  FILE REFERENCE: PO-16.2C1
;  CURRENT APPLICATION NUMBER: US/09/783,590
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 08/420,856
;  PRIOR FILING DATE: 1995-04-12
;  PRIOR APPLICATION NUMBER: 08/346,731
;  PRIOR FILING DATE: 1994-11-21
;  NUMBER OF SEQ ID NOS: 12485
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7099
;   LENGTH: 88
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (53)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-7099

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 88;
  Best Local Similarity   82.4%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 TTCAGGGAATCCGGGGG 20
          |||||||||  ||  |||
Db     19 TTCAGGGAAGCCCTGGG 35


RESULT 8
US-09-878-574-6486/c
; Sequence 6486, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 6486
;   LENGTH: 89
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701098483H1
US-09-878-574-6486

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 89;
  Best Local Similarity   82.4%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCGGG 18
          || || || ||||||||
Db     84 TTGTCTGGAAATCCGGG 68


RESULT 9
US-09-864-761-24917/c
; Sequence 24917, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
```

```
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24917
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL133163.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =  11
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =  14
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
;   OTHER INFORMATION: SWISSPROT HIT: Q9QUM9, EVALUE 9.00e-13
;   OTHER INFORMATION: EST_HUMAN HIT: BE797925.1, EVALUE 3.00e-35
;   OTHER INFORMATION: NT HIT: X59417.1, EVALUE 2.00e-35
US-09-864-761-24917

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 100;
  Best Local Similarity   82.4%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGG 17
          ||||||| |||| | ||
Db     55 TTTTTCAAGGAAGCTGG 39


RESULT 10
US-09-759-143-850/c
; Sequence 850, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 850
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR primer
US-09-759-143-850

  Query Match             60.0%;  Score 12;  DB 10;  Length 40;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          |||| || ||| | || |||
Db     36 TTTTACACGGATTTCGTGGG 17


RESULT 11
US-09-780-669-850/c
; Sequence 850, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 850
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR primer
US-09-780-669-850
```

```
  Query Match             60.0%;  Score 12;  DB 10;  Length 40;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TTTTTCAGGGAATCCGGGGG 20
           |||| || ||| | || |||
Db      36 TTTTACACGGATTTCGTGGG 17


RESULT 12
US-09-822-827-850/c
; Sequence 850, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C1
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 850
;   LENGTH: 40
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR primer
US-09-822-827-850

  Query Match             60.0%;  Score 12;  DB 10;  Length 40;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TTTTTCAGGGAATCCGGGGG 20
           |||| || ||| | || |||
Db      36 TTTTACACGGATTTCGTGGG 17


RESULT 13
US-09-864-761-32573
; Sequence 32573, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32573
;   LENGTH: 84
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000169.1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.76
;   OTHER INFORMATION: NT HIT: AF168106.1, EVALUE 3.70e-02
US-09-864-761-32573

  Query Match             60.0%;  Score 12;  DB 10;  Length 84;
  Best Local Similarity   75.0%;  Pred. No. 2.7e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TTTTTCAGGGAATCCGGGGG 20
           | ||| ||||| ||| | ||
Db      13 TGTTTGAGGGATTCCAGAGG 32


RESULT 14
US-09-864-761-32016
; Sequence 32016, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32016
;   LENGTH: 86
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AP000054.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.66
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.62
;   OTHER INFORMATION: SWISSPROT HIT: P17525, EVALUE 4.10e+00
;   OTHER INFORMATION: NT HIT: AF092047.1, EVALUE 2.80e+00
US-09-864-761-32016

  Query Match             60.0%;  Score 12;  DB 10;  Length 86;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TTTTTCAGGGAATCCGGGGG 20
           | ||| ||||| ||| | ||
Db      37 TGTTTGAGGGATTCCAGAGG 56


RESULT 15
US-09-736-169-2/c
; Sequence 2, Application US/09736169
; Patent No. US20010018055A1
; GENERAL INFORMATION:
;  APPLICANT: Briggs, Robert E.
;  APPLICANT:  Tatum, Fred M.
;  TITLE OF INVENTION: LKTA Deletion Mutant of P. haemolytica
;  FILE REFERENCE: 00295.75997
;  CURRENT APPLICATION NUMBER: US/09/736,169
;  CURRENT FILING DATE:  2000-12-15
;  PRIOR APPLICATION NUMBER: 09/280,852
;  PRIOR FILING DATE: 2000-07-03
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 26
;   TYPE: DNA
;   ORGANISM: Pasteurella cf. haemolytica
US-09-736-169-2

  Query Match             59.0%;  Score 11.8;  DB 10;  Length 26;
  Best Local Similarity   86.7%;  Pred. No. 2.9e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TTTCAGGGAATCCGG 17
           |||||| | ||||||
Db      15 TTTCAGCGGATCCGG 1


Search completed: November 23, 2002, 06:42:04
Job time : 18.25 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-5
Perfect score: 20
Sequence: 1 tttttcagggaatccggggg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.8	74.0	79	9	AI925202	AI925202 wm99h04.x
	2	14.8	74.0	96	14	F26680	F26680 HSPD14248 H
	3	13.8	69.0	49	9	AI756242	AI756242 EtESTea41
	4	13.8	69.0	64	14	H41671	H41671 yo06f01.s1
	5	13.6	68.0	43	17	AZ346730	AZ346730 1M0082A11
c	6	13.6	68.0	67	9	AI799398	AI799398 tw56a01.x
c	7	13.6	68.0	70	9	AI356377	AI356377 qz21d04.x
	8	13.6	68.0	85	9	AA854921	AA854921 aj78h10.s
	9	13.6	68.0	88	14	C01948	C01948 HUMGS000395
	10	13.6	68.0	90	10	AW102784	AW102784 xd38a06.x
c	11	13.6	68.0	100	12	BG672884	BG672884 DRNAAOH08
	12	13.4	67.0	63	14	T68724	T68724 yc29d01.s1
c	13	13.4	67.0	79	9	AA140294	AA140294 mr83c12.r
	14	13.2	66.0	50	9	AU107883	AU107883 AU107883
	15	13.2	66.0	67	9	AU014487	AU014487 AU014487
	16	13.2	66.0	73	14	R62624	R62624 yi12f02.s1
c	17	13.2	66.0	76	9	AA918583	AA918583 ol53a08.s
c	18	13.2	66.0	77	17	AZ769484	AZ769484 1M0570D15
	19	13.2	66.0	79	9	AA411778	AA411778 zu02f04.s
c	20	13.2	66.0	82	12	BE802294	BE802294 sr28a12.y
c	21	13.2	66.0	97	9	AI649260	AI649260 uk27d06.x
c	22	13.2	66.0	100	13	BJ534085	BJ534085 BJ534085
	23	13	65.0	91	17	AZ623789	AZ623789 1M0461B19
	24	12.8	64.0	37	9	AA992758	AA992758 ot84f03.s
c	25	12.8	64.0	50	9	AU107657	AU107657 AU107657
c	26	12.8	64.0	76	14	T86922	T86922 yd86h10.s1
	27	12.8	64.0	91	14	T91269	T91269 yd60g11.s1
c	28	12.8	64.0	94	13	BM507198	BM507198 ih25d04.y
c	29	12.6	63.0	34	9	AI632285	AI632285 tt22b07.x
c	30	12.6	63.0	41	9	AU269153	AU269153 AU269153
c	31	12.6	63.0	46	10	AV671661	AV671661 AV671661
c	32	12.6	63.0	53	17	B01982	B01982 cSRL-144H11
c	33	12.6	63.0	57	9	AU259067	AU259067 AU259067
	34	12.6	63.0	69	17	BH625831	BH625831 1007109F0
	35	12.6	63.0	77	17	AL759161	AL759161 Arabidops
	36	12.6	63.0	85	14	BQ821331	BQ821331 1030091E0
	37	12.6	63.0	87	9	AI173945	AI173945 uf05h07.x
c	38	12.6	63.0	89	10	AW248154	AW248154 2819681.5
c	39	12.6	63.0	90	10	AW549193	AW549193 L0047E04-
	40	12.6	63.0	94	9	AI188318	AI188318 qd13b02.x
	41	12.6	63.0	94	17	AZ424250	AZ424250 1M0203F18
	42	12.6	63.0	94	17	BH642503	BH642503 1008041A0
	43	12.6	63.0	95	17	AZ853253	AZ853253 2M0156P13
c	44	12.6	63.0	97	9	AA864216	AA864216 oh39d02.s
c	45	12.6	63.0	97	13	BM513037	BM513037 kx83g10.y

ALIGNMENTS

RESULT 1
AI925202
LOCUS AI925202 79 bp mRNA linear EST 07-MAR-2000
DEFINITION wm99h04.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2444119 3' similar to SW:BAT2_HUMAN P48634 LARGE PROLINE-RICH PROTEIN BAT2 ;, mRNA sequence.
ACCESSION AI925202
VERSION AI925202.1 GI:5661166
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 79)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```
          Trace considered overall poor quality
          Insert Length: 768   Std Error: 0.00
          Seq'primer: -40UP from Gibco
          High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2444119"
                     /clone_lib="NCI_CGAP_Ut2"
                     /tissue_type="moderately-differentiated endometrial
                     adenocarcinoma, 3 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.85 kb. Life Technologies catalog #:
                     11539-012"
BASE COUNT       11 a     26 c     38 g      4 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 9;  Length 79;
  Best Local Similarity   88.9%;  Pred. No. 9.1e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TTTCAGGGAATCCGGGGG 20
           ||| |||||| |||||||
Db      10 TTTTAGGGAACCCGGGGG 27


RESULT 2
F26680
LOCUS       F26680                    96 bp    mRNA    linear   EST 13-MAY-1999
DEFINITION  HSPD14248 HM3 Homo sapiens cDNA clone s4000058G04, mRNA sequence.
ACCESSION   F26680
VERSION     F26680.1  GI:4812306
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Lanfranchi,G., Muraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,
            Pandolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.
  TITLE     Identification of 4370 expressed sequence tags from a
            3'-end-specific cDNA library of human skeletal muscle by DNA
            sequencing and filter hybridization
  JOURNAL   Genome Res. 6 (1), 35-42 (1996)
  MEDLINE   96276048
COMMENT     Contact: Valle G.
            CRIBI Biotechnology Centre
            University of Padua
            Via Trieste 75, 35121 Padua, Italy
            ABI Chromatograms and other information are available on WWW at
            http://grup.bio.unipd.it.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="s4000058G04"
                     /clone_lib="HM3"
                     /sex="female"
                     /tissue_type="pectoral muscle (after mastectomy)"
                     /note="Vector: pcDNAII (Invitrogen); Site_1: BstXI;
                     Site_2: NotI; The library was constructed by G.
                     Lanfranchi. This library is not subtracted nor normalized.
                     The first strand cDNA was primed with a biotinylated
                     oligo-dT-NotI primer
                     (5'-biotin-AACCCGGCTCGAGCGGCCGCTTTTTTTTTTTTTTTTT-3'). The
                     ds cDNA was sonicated and size-selected in the range
                     350-550  bp. The 3' specific fragments were selected by
                     streptavidin coated magnetic beads, ligated to
                     non-palindromic BstXI adapters, NotI digested and
                     directionally cloned into BstXI-NotI cut pcDNAII vector."
BASE COUNT       19 a     25 c     35 g     17 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 14;  Length 96;
  Best Local Similarity   88.9%;  Pred. No. 9.1e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 TTTTCAGGGAATCCGGGG 19
           |||||||||  |||||||
Db      79 TTTTCAGGGCCTCCGGGG 96


RESULT 3
AI756242
LOCUS       AI756242                  49 bp    mRNA    linear   EST 18-JAN-2000
DEFINITION  EtESTea41e11.y1 Eimeria S5-2 Sporozoite stage Eimeria tenella cDNA
            5' similar to SW:U2AG_HUMAN Q01081 SPLICING FACTOR U2AF 35 KD
            SUBUNIT ;, mRNA sequence.
ACCESSION   AI756242
VERSION     AI756242.1  GI:5149965
KEYWORDS    EST.
SOURCE      Eimeria tenella.
  ORGANISM  Eimeria tenella
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
            Eimeria.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Liberator,P., Diaz,C., Tang,K., Marra,M., Hillier,L., Kucaba,T.,
            Martin,J., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen
            ,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey
            ,N., Schurk,R., Ritter,E., Kohn,S., Florence,N., Shin,T., Jackson
            ,Y., Cardenas,M., McCann,R., Waterston,R., Wilson,R. and Sibley,D.
  TITLE     WashU-Merck Eimeria tenella project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Sibley, Ph.D.
            WashU-Merck Eimeria tenella project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Contact David Sibley (toxoest@borcim.wustl.edu) for further
            information relating to organism, libraries, or clone availability.
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40RP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Eimeria tenella"
                     /strain="LS18"
                     /db_xref="taxon:5802"
                     /clone_lib="Eimeria S5-2 Sporozoite stage"
                     /dev_stage="Sporozoite"
                     /lab_host="SOLR E. coli"
                     /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI
                     ; Sporozoites were obtained from in vitro sporulated and
                     excysted oocysts of E. tenella grown in chickens.  cDNA
                     was synthesized from poly mRNA using an oligo-dT primer
                     containing a XhoI site. Following second strand synthesis,
                     EcoRI adapters were ligated to  the cDNA and products were
                     size-selected on Sephacryl S500.  cDNAs were digested with
                     EcoRI/XhoI and cloned into lambda Zap II (Stratagene).
                     Clones were converted to phagemids by mass excision  using
                     ExAssist helper phage and SOLR cells (Stratagene ).
                     Insert  sizes range from 1.2-2.9 kb."
BASE COUNT        8 a     13 c     15 g     13 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 49;
  Best Local Similarity   88.2%;  Pred. No. 2.7e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TTTCAGGGAATCCGGGG 19
           |||||||||| |||| |
```

```
Db      15 TTTCAGGGAAGCCGGTG 31


RESULT 4
H41671
LOCUS       H41671                    64 bp    mRNA    linear   EST 31-JUL-1995
DEFINITION  yo06f01.s1 Soares adult brain N2b5HB55Y Homo sapiens cDNA clone
            IMAGE:177145 3' similar to SP:RS9_RAT P29314 40S RIBOSOMAL PROTEIN
            ;, mRNA sequence.
ACCESSION   H41671
VERSION     H41671.1  GI:917723
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1652
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 1652   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Homo sapiens"
                     /db_xref="GDB:3839341"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:177145"
                     /clone_lib="Soares adult brain N2b5HB55Y"
                     /sex="Male"
                     /dev_stage="55-year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: brain; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 53. Library constructed by Bento
                     Soáres and M.Fatima Bonaldo. The adult brain RNA was
                     provided by Dr. Donald H. Gilden. Tissue was acquired
                     17-18 hours after death which occurred in consequence of a
                     ruptured aortic aneurysm. RNA was prepared from a pool of
                     tissues representing the following areas of the brain:
                     frontal, parietal, temporal and occipital cortex from the
                     left and right hemispheres, subcortical white matter,
                     basal ganglia, thalamus, cerebellum, midbrain, pons and
                     medulla."
BASE COUNT       13 a     15 c     20 g     14 t      2 others
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 14;  Length 64;
  Best Local Similarity   88.2%;  Pred. No. 2.7e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTCAGGGAATCCGGGGG 20
           ||| |||||||| ||||
Db       6 TTCTGGGAATCCAGGGG 22


RESULT 5
AZ346730
LOCUS       AZ346730                  43 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0082A11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0082A11 F, DNA sequence.
ACCESSION   AZ346730
VERSION     AZ346730.1  GI:10425967
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0082  row: A  column: 11
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 43.
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0082A11"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        7 a     13 c      9 g     14 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 43;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
```

```
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          ||||   |||||||| ||||
Db      4 TTTTCTGGGGAATCCAGGGG 23


RESULT 6
AI799398/c
LOCUS       AI799398                  67 bp    mRNA    linear   EST 15-DEC-1999
DEFINITION  tw56a01.x1 NCI_CGAP_Ut3 Homo sapiens cDNA clone IMAGE:2263656 3'
            similar to SW:PRPM_HUMAN P10161 SALIVARY PROLINE-RICH PROTEIN PO
            ;contains element MSR1 repetitive element ;, mRNA sequence.
ACCESSION   AI799398
VERSION     AI799398.1  GI:5364870
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1905   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2263656"
                     /clone_lib="NCI_CGAP_Ut3"
                     /tissue_type="poorly-differentiated endometrial
                     adenocarcinoma, 2 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.45 kb. Life Technologies catalog #:
                     11541-018"
BASE COUNT       20 a     22 c     25 g      0 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 67;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          |||||| |||   |||||||
Db     50 TTTTTCCGGGGCCCCGGGGG 31


RESULT 7
AI356377/c
LOCUS       AI356377                  70 bp    mRNA    linear   EST 16-FEB-1999
DEFINITION  qz21d04.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2027527 3'
            similar to TR:Q64306 Q64306 PROLINE-RICH PROTEIN. ;contains element
            MSR1 repetitive element ;, mRNA sequence.
ACCESSION   AI356377
VERSION     AI356377.1  GI:4107998
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
            M.D., Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 1100   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2027527"
                     /clone_lib="NCI_CGAP_CLL1"
                     /tissue_type="B-cell, chronic lymphotic leukemia"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTGCTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT       21 a     27 c     13 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 70;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          |||| |  ||||| ||||||
Db     51 TTTTCCCCGGAATTCGGGGG 32


RESULT 8
AA854921
LOCUS       AA854921                  85 bp    mRNA    linear   EST 04-JAN-1999
DEFINITION  aj78h10.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
            IMAGE:1402627 3' similar to TR:Q92801 Q92801 HYPOTHETICAL 28.3 KD
            PROTEIN. ;, mRNA sequence.
ACCESSION   AA854921
VERSION     AA854921.1  GI:2942459
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
            , Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 1001   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .85
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1402627"
                      /clone_lib="Soares_parathyroid_tumor_NbHPA"
                      /tissue_type="parathyroid tumor"
                      /dev_stage="adult"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
                      ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                      RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                      primer
                      [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTTT
                      TTTTT-3'], double-stranded cDNA was size selected, ligated
                      to Eco RI adapters (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of a modified pT7T3
                      vector (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by Bento
                      Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
                      adenomas was kindly provided by Dr. Stephen Marx, National
                      Institute of Diabetes and Digestive and Kidney Diseases,
                      NIH."
BASE COUNT       14 a     19 c     25 g     27 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 85;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          |||||| |||||| ||| |
Db     30 TTTTTCTCGGAATCTGGGCG 49


RESULT 9
C01948
LOCUS       C01948                    88 bp    mRNA    linear   EST 23-JUL-1996
DEFINITION  HUMGS0003954 Human adult (K.Okubo) Homo sapiens cDNA, mRNA
            sequence.
ACCESSION   C01948
VERSION     C01948.1  GI:1434178
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Okubo,K.
  TITLE     BodyMap; human gene expression database
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Okubo,K.
            Institute for Molecular and Cellular Biol
            Osaka University
            1-3,Yamada-oka, Suita, Osaka Pref. 565, Japan
            Tel: 06-877-5111(ex.3315)
            Email: kousaku@imcb.osaka-u.ac.jp
            Human Gene Signature, 3'-directed cDNA sequence. We are not
            submitting the same cDNA sequence redundantly to DDBJ since 1993.
            For the abundance information of clones with this sequence in this
            library and as well as in other 3'-directed libraries, see '
            http://www.imcb.osaka-u.ac.jp/bodymap'. The sequences of the clones
            represented by this GS sequences is also found there.
FEATURES              Location/Qualifiers
     source           1. .88
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="Human adult (K.Okubo)"
                      /dev_stage="adult"
                      /note="Organ: blood; Vector: l-gt-11; Site_1: Eco-RI;
                      Monocytes were prepared from blood by ficoll-hypaque,
                      percoll  and T cell rosetting purification steps (purity:
                      96 %). mRNA was prepared from  activated monocytes from a
                      patient with rheumatoid arthritis. mRNA was reverse
                      transcribed with MuLV. Using Eco-RI linkers cDNA was
                      cloned into l-gt-11  vector arms. The cDNA library was
                      screened by differential hybidization  using radioactively
                      marked ss-cDNA from activated and non-activated
                      monocytes.    "
BASE COUNT       28 a     15 c     12 g     33 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 88;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          ||||| |||||||||   ||
Db     25 TTTTTAAGGGAATCCACTGG 44


RESULT 10
AW102784
LOCUS       AW102784                  90 bp    mRNA    linear   EST 19-OCT-1999
DEFINITION  xd38a06.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2596018 3',
            mRNA sequence.
ACCESSION   AW102784
VERSION     AW102784.1  GI:6073397
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 73.
FEATURES              Location/Qualifiers
     source           1. .90
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2596018"
                      /clone_lib="NCI_CGAP_Ov23"
                      /tissue_type="tumor, 5 pooled (see description)"
                      /lab_host="DH10B"
                      /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 1.35 kb. Tumor types include: mixed
                      Mullerian tumor, papillary serous, clear cell, spindle
```

```
                cell.  All are primary tumors, metastasis positive. Life
                Technologies catalog #: 11534-013"
BASE COUNT      17 a     16 c     18 g     39 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 10;  Length 90;
  Best Local Similarity  80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          ||||| | |||| ||| |||
Db     38 TTTTTAAAGGAAACCGAGGG 57


RESULT 11
BG672884/c
LOCUS       BG672884                 100 bp    mRNA    linear   EST 30-APR-2001
DEFINITION  DRNAAOH08 Rat DRG Library Rattus norvegicus cDNA clone DRNAAOH08
            5', mRNA sequence.
ACCESSION   BG672884
VERSION     BG672884.1  GI:13894983
KEYWORDS    EST.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Xiao,H.S., Huang,Q.H., Zhang,F.X., Bao,L., Lu,Y.J., Guo,C., Yang,L.
            , Huang,W.J., Fu,G., Xu,S.H., Cheng,X.P., Yan,Q., Zhu,Z.D., Zhang
            ,X., Chen,Z., Han,Z.G. and Zhang,X.
  TITLE     Identification of gene expression profile of dorsal root ganglion
            in the rat peripheral axotomy model of neuropathic pain
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (12), 8360-8366 (2002)
  MEDLINE   22056133
COMMENT     Contact: Zhang Xu
            Laboratory of Sensory System
            Institute of Neuroscience
            320 Yue Yang Road, Shanghai 200031, P.R.China
            Tel: 86-21-64748700-121
            Fax: 86-21-64713446
            Email: xu.zhang@ion.ac.cn
            This clone is also available at Chinese National Human Genome
            Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park,
            Pudong New Area, P.R.China. Please contact with Zhang Xu
            (xu.zhang@ion.ac.cn) or Han Zeguang (hanzg@chgc.sh.cn)
            PCR PRimers
            FORWARD: T3
            BACKWARD: T7
            Seq primer: T3
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .100
                      /organism="Rattus norvegicus"
                      /strain="Sprague-Dawley"
                      /db_xref="taxon:10116"
                      /clone="DRNAAOH08"
                      /clone_lib="Rat DRG Library"
                      /sex="male"
                      /tissue_type="dorsal root ganglion"
                      /dev_stage="adult"
                      /note="Total RNA was isolated from hypothalamus and
                      transcripted into cDNA,  which was then used as templat in
                      PCR.The PCR products were separated  on sequencing gel.
                      The differential bands were cut, reamplified, cloned into
                      pMD18-T vector and confirmed by Northern blot."
BASE COUNT      38 a     16 c     17 g     28 t      1 others
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 12;  Length 100;
  Best Local Similarity  80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCCGGGGG 20
          ||||| | |||| ||| |||
Db     95 TTTTTTAGGGAAACCGAGGG 76


RESULT 12
T68724
LOCUS       T68724                   63 bp    mRNA    linear   EST 22-FEB-1995
DEFINITION  yc29d01.s1 Stratagene liver (#937224) Homo sapiens cDNA clone
            IMAGE:82081 3' similar to gb:X02162 APOLIPOPROTEIN A-I PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   T68724
VERSION     T68724.1  GI:679872
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence starts: 1
            High qality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Seq primer: -21m13
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .63
                      /organism="Homo sapiens"
                      /db_xref="GDB:499138"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:82081"
                      /clone_lib="Stratagene liver (#937224)"
                      /sex="male"
                      /dev_stage="49 years old"
                      /lab_host="SOLR cells (kanamycin resistant)"
                      /note="Organ: liver; Vector: pBluescript SK; Site_1: EcoRI
                      ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                      dT. Hepatectomy from normal male caucasian. Average insert
                      size: 1.1 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                      GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                      CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT      8 a     18 c     11 g     19 t      7 others
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 14;  Length 63;
  Best Local Similarity  93.3%;  Pred. No. 4.2e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TTTTTCAGGGAATCC 15
          ||||||||||| ||||
Db     36 TTTTTCAGGGTATCC 50


RESULT 13
AA140294/c
LOCUS       AA140294                 79 bp    mRNA    linear   EST 11-FEB-1997
```

```
DEFINITION  mr83c12.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
            IMAGE:604054 5' similar to SW:IDHG_RAT P41565 ISOCITRATE
            DEHYDROGENASE ;, mRNA sequence.
ACCESSION   AA140294
VERSION     AA140294.1  GI:1702550
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:369486
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Mus musculus"
                     /strain="NIH Swiss"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:604054"
                     /clone_lib="Stratagene mouse heart (#937316)"
                     /sex="pooled"
                     /tissue_type="heart"
                     /dev_stage="13 day embryos"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: heart; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. 93 pooled NIH/Swiss 13 day embryo hearts.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       20 a     16 c     20 g     23 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 79;
  Best Local Similarity   93.3%;  Pred. No. 4.2e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 TTCAGGGAATCCGGG 18
          ||||||||| |||||
Db     59 TTCAGGGAACCCGGG 45


RESULT 14
AU107883
LOCUS       AU107883                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU107883 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            ZRV61130, mRNA sequence.
ACCESSION   AU107883
VERSION     AU107883.1  GI:13557405
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ZRV61130"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        8 a     13 c     12 g     16 t      1 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 50;
  Best Local Similarity   83.3%;  Pred. No. 5.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TTTTCAGGGAATCCGGGG 19
            || ||||| || ||||||
Db     32 TTCTCAGGCAAGCCGGGG 49


RESULT 15
AU014487
LOCUS       AU014487                 67 bp    mRNA    linear   EST 03-AUG-1998
DEFINITION  AU014487 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc09929, mRNA sequence.
ACCESSION   AU014487
VERSION     AU014487.1  GI:3369278
KEYWORDS    EST.
SOURCE      fission yeast.
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Schizosaccharomyces pombe"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc09929"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /sex="h minus"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
```

```
BASE COUNT       27 a      1 c     11 g     28 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 67;
  Best Local Similarity   83.3%;  Pred. No. 5.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TTTTTCAGGGAATCCGGG 18
           ||||||||| |||  |||
Db      10 TTTTTCAGGAAATATGGG 27


Search completed: November 26, 2002, 04:07:03
Job time : 768.8 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-6
Perfect score: 20
Sequence: 1 gggtagttcaggcgggagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	14.2	71.0	51	6	AX158137	AX158137 Sequence
2	14.2	71.0	75	1	ECCTRNAV	X70451 E.coli gene
c 3	14.2	71.0	95	6	AX088789	AX088789 Sequence
4	14.2	71.0	96	6	AX088790	AX088790 Sequence
5	13.6	68.0	63	6	E27556	E27556 Isolation a
6	13.6	68.0	76	1	ECOTRNAPHE	M24882 E. coli tra
7	13.4	67.0	42	6	AR104424	AR104424 Sequence
8	13.4	67.0	63	6	AR127588	AR127588 Sequence
9	13.4	67.0	63	6	I75257	I75257 Sequence 6
10	13.4	67.0	73	6	I13478	I13478 Sequence 12
11	13.2	66.0	28	6	I67871	I67871 Sequence 7
c 12	13.2	66.0	51	6	AX199369	AX199369 Sequence
c 13	13.2	66.0	51	6	AX199370	AX199370 Sequence
c 14	12.8	64.0	21	6	AR148748	AR148748 Sequence
c 15	12.8	64.0	40	6	AX025223	AX025223 Sequence
c 16	12.8	64.0	48	6	AR031652	AR031652 Sequence
c 17	12.8	64.0	48	6	I90270	I90270 Sequence 11
18	12.8	64.0	76	1	ECOTRN	K00164 E.coli Asn-
19	12.8	64.0	77	1	ECOTRN1	M35841 E.coli Asn-
20	12.8	64.0	84	1	ECOTRN2	M35840 E.coli Asn-
21	12.8	64.0	87	1	ECOTRN3A	M24867 E.coli Asn-
22	12.8	64.0	87	1	ECOTRN3B	M35839 E.coli Asn-
c 23	12.6	63.0	32	6	AX026032	AX026032 Sequence
24	12.6	63.0	36	6	AX452383	AX452383 Sequence
25	12.6	63.0	40	9	HSZ74584	Z74584 H.sapiens j
c 26	12.6	63.0	45	6	AX417920	AX417920 Sequence
c 27	12.6	63.0	48	6	AX014267	AX014267 Sequence
c 28	12.6	63.0	51	6	AX003711	AX003711 Sequence
c 29	12.6	63.0	51	6	AX158138	AX158138 Sequence
c 30	12.6	63.0	51	6	AX256314	AX256314 Sequence
c 31	12.6	63.0	60	6	AX100154	AX100154 Sequence
c 32	12.6	63.0	60	6	AX100155	AX100155 Sequence
c 33	12.6	63.0	60	6	AX100156	AX100156 Sequence
c 34	12.6	63.0	60	6	AX100157	AX100157 Sequence
c 35	12.6	63.0	63	6	AX190431	AX190431 Sequence
c 36	12.6	63.0	64	6	AX003715	AX003715 Sequence
c 37	12.6	63.0	64	6	AX023351	AX023351 Sequence
c 38	12.6	63.0	68	6	AX100153	AX100153 Sequence
c 39	12.6	63.0	75	6	AX353451	AX353451 Sequence
40	12.6	63.0	76	1	BACTRF	K00332 B.stearothe
41	12.6	63.0	79	6	AX003726	AX003726 Sequence
c 42	12.6	63.0	79	6	AX003727	AX003727 Sequence
43	12.6	63.0	80	6	AX018524	AX018524 Sequence
44	12.6	63.0	80	6	AX018528	AX018528 Sequence
45	12.6	63.0	80	6	AX018555	AX018555 Sequence

ALIGNMENTS

RESULT 1
AX158137/c
LOCUS AX158137 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 1465 from Patent WO0140521.
ACCESSION AX158137
VERSION AX158137.1 GI:14539468
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 51)
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof

```
  JOURNAL   Patent: WO 0140521-A 1465 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (1466 is other entry)
                     Accession number cg29363109"
BASE COUNT       10 a     19 c     10 g     12 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 2.7e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGGTAGTTCAGGCGGGAGC 19
          |||| || ||| |||||||
Db     28 GGGTGGTACAGCCGGGAGC 10


RESULT 2
ECCTRNAV
LOCUS       ECCTRNAV                  75 bp    mRNA    linear   BCT 16-JUN-1993
DEFINITION  E.coli gene for chimeric transfer RNA-Val.
ACCESSION   X70451
VERSION     X70451.1  GI:312130
KEYWORDS    chimeric; transfer RNA-Val.
SOURCE      Escherichia coli.
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Meinnel,T., Mechulam,Y., Lazennec,C., Blanquet,S. and Fayat,G.
  TITLE     Critical role of the acceptor stem of tRNAs(Met) in their
            aminoacylation by Escherichia coli methionyl-tRNA synthetase
  JOURNAL   J. Mol. Biol. 229 (1), 26-36 (1993)
  MEDLINE   93132797
   PUBMED   8421312
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Escherichia coli"
                     /db_xref="taxon:562"
     tRNA            1. .75
                     /product="tRNA-Val"
BASE COUNT       12 a     26 c     26 g      9 t      2 others
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 1;  Length 75;
  Best Local Similarity   80.0%;  Pred. No. 2.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GGGTAGTTCAGGCGGGAGCG 20
          |||||| ||||  ||||| |
Db      5 GGGTAGCTCAGCNGGGAGAG 24


RESULT 3
AX088789/c
LOCUS       AX088789                  95 bp    DNA     linear   PAT 17-MAR-2001
DEFINITION  Sequence 115 from Patent WO0114416.
ACCESSION   AX088789
VERSION     AX088789.1  GI:13397585
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Neeper,M.P., Mcclements,W.L., Jansen,K.U., Schultz,L.D., Chen,L.
            and Wang,X.M.
  TITLE     Synthetic human papillomavirus genes
  JOURNAL   Patent: WO 0114416-A 115 01-MAR-2001;
            Merck & Co., Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Codon-Optimized HPV6 E2 fragment"
BASE COUNT       23 a     36 c     23 g     13 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 95;
  Best Local Similarity   84.2%;  Pred. No. 2.6e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGTAGTTCAGGCGGGAGCG 20
          | | ||||||||||| ||||
Db     92 GATCGTTCAGGCGGTAGCG 74


RESULT 4
AX088790
LOCUS       AX088790                  96 bp    DNA     linear   PAT 17-MAR-2001
DEFINITION  Sequence 116 from Patent WO0114416.
ACCESSION   AX088790
VERSION     AX088790.1  GI:13397586
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Neeper,M.P., Mcclements,W.L., Jansen,K.U., Schultz,L.D., Chen,L.
            and Wang,X.M.
  TITLE     Synthetic human papillomavirus genes
  JOURNAL   Patent: WO 0114416-A 116 01-MAR-2001;
            Merck & Co., Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Codon-Optimized HPV6 E2 fragment"
BASE COUNT       13 a     23 c     40 g     20 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 96;
  Best Local Similarity   84.2%;  Pred. No. 2.6e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGTAGTTCAGGCGGGAGCG 20
          | | ||||||||||| ||||
Db     73 GATCGTTCAGGCGGTAGCG 91


RESULT 5
E27556
LOCUS       E27556                    63 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Isolation and cloning of nucleic acid molecule by using hairpin
            loop nucleic acid probe molecule.
ACCESSION   E27556
VERSION     E27556.1  GI:13018221
KEYWORDS    JP 1999206381-A/4.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Jun,F. and Koji,S.
  TITLE     Isolation and cloning of nucleic acid molecule by using hairpin
            loop nucleic acid probe molecule
  JOURNAL   Patent: JP 1999206381-A 4 03-AUG-1999;
            AISIN COSMOS R&D CO LTD
COMMENT     OS   Unidentified
            PN   JP 1999206381-A/4
            PD   03-AUG-1999
            PF   29-JAN-1998 JP 1998016987
            PR
            PI   JUN FUJIWARA,KOJI SHIGEMORI
```

```
            PC   C12N15/09,C12Q1/68,C12N15/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .63
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        4 a     20 c     25 g     14 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 63;
  Best Local Similarity  80.0%;  Pred. No. 5.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GGGTAGTTCAGGCGGGAGCG 20
          ||| || ||||||||  |||
Db     30 GGGCAGGTCAGGCGGCCGCG 49


RESULT 6
ECOTRNAPHE
LOCUS       ECOTRNAPHE                76 bp    tRNA    linear   BCT 15-APR-1994
DEFINITION  E. coli transfer RNA-Phe.
ACCESSION   M24882
VERSION     M24882.1  GI:290481
KEYWORDS    transfer RNA-Phe.
SOURCE      Escherichia coli tRNA.
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Kitchingman,G.R. and Fournier,M.J.
  TITLE     Modification-deficient transfer ribonucleic acids from relaxed
            control Escherichia coli: structures of the major undermodified
            phenylalanine and leucine transfer RNAs produced during leucine
            starvation
  JOURNAL   Biochemistry 16 (10), 2213-2220 (1977)
  MEDLINE   77182103
   PUBMED   324516
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Escherichia coli"
                     /db_xref="taxon:562"
     tRNA            1. .76
                     /product="tRNA-Phe"
                     /note="codon recognized: UUC"
BASE COUNT       15 a     21 c     24 g     13 t      3 others
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 1;  Length 76;
  Best Local Similarity  75.0%;  Pred. No. 5.2e+04;
  Matches   15;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGGTAGTTCAGGCGGGAGCG 20
          || ||| ||||:|||:|| |
Db      5 GGATAGCTCAGDCGGDAGAG 24


RESULT 7
AR104424
LOCUS       AR104424                  42 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 9 from patent US 6093573.
ACCESSION   AR104424
VERSION     AR104424.1  GI:12817132
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Beamer,L.J., Carroll,S.F. and Eisenberg,D.
  TITLE     Three-dimensional structure of bactericidal/permeability-increasing
            protein (BPI)
  JOURNAL   Patent: US 6093573-A 9 25-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="unknown"
BASE COUNT       11 a      8 c     18 g      5 t
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 6;  Length 42;
  Best Local Similarity  93.3%;  Pred. No. 6.8e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GTTCAGGCGGGAGCG 20
          ||||||||||| ||
Db      3 GTTCAGGCGGGACCG 17


RESULT 8
AR127588
LOCUS       AR127588                  63 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 6 from patent US 6180774.
ACCESSION   AR127588
VERSION     AR127588.1  GI:14114183
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Brown,S.Marie., Dean,D.Allen., Fromm,M.Ernest. and
            Sanders,P.Rigden.
  TITLE     Synthetic DNA sequences having enhanced expression in
            monocotyledonous plants and method for preparation thereof
  JOURNAL   Patent: US 6180774-A 6 30-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="unknown"
BASE COUNT       14 a     20 c     15 g     14 t
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 6;  Length 63;
  Best Local Similarity  93.3%;  Pred. No. 6.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 AGTTCAGGCGGGAGC 19
          ||||||||||  ||||
Db     26 AGTTCAGGCGCGAGC 40


RESULT 9
I75257
LOCUS       I75257                    63 bp    DNA     linear   PAT 03-APR-1998
DEFINITION  Sequence 6 from patent US 5689052.
ACCESSION   I75257
VERSION     I75257.1  GI:3011398
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Brown,S.Marie., Dean,D.Allen., Fromm,M.Ernest. and
            Sanders,P.Rigden.
  TITLE     Synthetic DNA sequences having enhanced expression in
            monocotyledonous plants and method for preparation thereof
  JOURNAL   Patent: US 5689052-A 6 18-NOV-1997;
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="unknown"
BASE COUNT       14 a     20 c     15 g     14 t
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 6;  Length 63;
  Best Local Similarity  93.3%;  Pred. No. 6.6e+04;
```

```
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 AGTTCAGGCGGGAGC 19
          ||||||||||  ||||
Db     26 AGTTCAGGCGCGAGC 40


RESULT 10
I13478
LOCUS       I13478                    73 bp    DNA     linear   PAT 26-JUL-1995
DEFINITION  Sequence 12 from patent US 5436391.
ACCESSION   I13478
VERSION     I13478.1  GI:910819
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Fujimoto,H., Ito,K., Yamamoto,M. and Shimamoto,K.
  TITLE     Synthetic insecticidal gene, plants of the genus oryza transformed
            with the gene, and production thereof
  JOURNAL   Patent: US 5436391-A 12 25-JUL-1995;
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="unknown"
BASE COUNT       15 a     25 c     21 g     12 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 73;
  Best Local Similarity   93.3%;  Pred. No. 6.5e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 AGTTCAGGCGGGAGC 19
          ||||||||  ||||||
Db     29 AGTTCAGGAGGGAGC 43


RESULT 11
I67871
LOCUS       I67871                    28 bp    DNA     linear   PAT 04-FEB-1998
DEFINITION  Sequence 7 from patent US 5674683.
ACCESSION   I67871
VERSION     I67871.1  GI:2829993
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Kool,E.T.
  TITLE     Stem-loop and circular oligonucleotides and method of using
  JOURNAL   Patent: US 5674683-A 7 07-OCT-1997;
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unknown"
BASE COUNT        4 a      4 c     16 g      4 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 28;
  Best Local Similarity   83.3%;  Pred. No. 8.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GGGTAGTTCAGGCGGGAG 18
          ||||  ||  ||||  |||||
Db      6 GGGTGGTCCAGGAGGGAG 23


RESULT 12
AX199369/c
LOCUS       AX199369                  51 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 299 from Patent WO0151670.
ACCESSION   AX199369
VERSION     AX199369.1  GI:15389759
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.D.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0151670-A 299 19-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (300 is other entry)
                     Accession number cg43973526"
BASE COUNT       10 a     17 c     16 g      8 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 8.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGTAGTTCAGGCGGGAGC 19
          || | | ||||||||||
Db     19 GGGATTACAGGCGGGAGC 2


RESULT 13
AX199370/c
LOCUS       AX199370                  51 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 300 from Patent WO0151670.
ACCESSION   AX199370
VERSION     AX199370.1  GI:15389761
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.D.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0151670-A 300 19-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (299 is other entry)
                     Accession number cg43973526"
BASE COUNT       10 a     16 c     16 g      9 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 8.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGTAGTTCAGGCGGGAGC 19
          || | | ||||||||||
Db     19 GGGATTACAGGCGGGAGC 2


RESULT 14
AR148748/c
LOCUS       AR148748                  21 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 105 from patent US 6225451.
ACCESSION   AR148748
VERSION     AR148748.1  GI:15112838
KEYWORDS    .
SOURCE      Unknown.
```

```
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Ballinger,D.G., Ding,W., Wagner,S. and Hess,M.A.
  TITLE     Chromosome 11-linked coronary heart disease susceptibility gene
            CHD1
  JOURNAL   Patent: US 6225451-A 105 01-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        3 a      8 c      4 g      6 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GTAGTTCAGGCGGGAG 18
           ||||| ||||| ||||
Db      19 GTAGTCCAGGCAGGAG 4


RESULT 15
AX025223/c
LOCUS       AX025223                 40 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 6 from Patent DE19849348.
ACCESSION   AX025223
VERSION     AX025223.1  GI:10186937
KEYWORDS    .
SOURCE      rhesus monkey.
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Schmidt,M. and Kalle,C.V.
  JOURNAL   Patent: DE 19849348-A 6 27-APR-2000;
            UNIV LUDWIGS ALBERT (DE)
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Macaca mulatta"
                     /db_xref="taxon:9544"
     LTR             31. .40
                     /note="Die ersten 10 Nucleotide von pLN 5'LTR
                     (NCBI/GenBank: M28245)"
BASE COUNT        9 a     13 c      7 g      9 t      2 others
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 40;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGG 16
           ||||  ||||||||||
Db      40 GGGTCTTTCAGGCGGG 25


Search completed: December  3, 2002, 18:12:34
Job time : 309.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-6
Perfect score: 20
Sequence: 1 gggtagttcaggcgggagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31436	Human neuropilin m
c	2	18	90.0	30	22	AAF83316	Human neuropilin-1
c	3	14.2	71.0	51	22	AAI74524	Human silent SNP c
c	4	14.2	71.0	65	24	ABN52875	Mouse spliced tran
c	5	14.2	71.0	95	22	AAF75497	Codon-optimised HP
	6	14.2	71.0	96	22	AAF75498	Codon-optimised HP
	7	13.6	68.0	30	24	ABK53798	DMS:acceptor oxido
	8	13.6	68.0	63	20	AAZ09153	Hairpin type nucle
	9	13.6	68.0	76	24	ABK87514	Human immunodefici
	10	13.6	68.0	76	24	ABK87515	Human immunodefici
	11	13.6	68.0	76	24	ABK87516	Human immunodefici
	12	13.6	68.0	76	24	ABK87517	Human immunodefici
	13	13.6	68.0	76	24	ABK87518	Human immunodefici
	14	13.6	68.0	76	24	ABK87519	Human immunodefici
c	15	13.6	68.0	88	21	AAC09858	Human secreted pro
	16	13.4	67.0	63	19	AAV00294	Bacillus thuringie
	17	13.4	67.0	63	22	AAF73212	Oligonucleotide #3
	18	13.2	66.0	28	17	AAT39026	Circular oligo com
	19	13.2	66.0	28	18	AAT94718	Target sequence bi
c	20	13.2	66.0	33	16	AAT05187	Mutant HSV-1 thymi
c	21	13.2	66.0	33	20	AAX57634	Codons 165-175 of
c	22	13.2	66.0	33	20	AAX15363	Codons 165-175 of
	23	13.2	66.0	37	21	AAA28282	Human beta-secreta
c	24	13.2	66.0	51	22	AAH89518	Human coding sequ
c	25	13.2	66.0	51	22	AAH89519	Human coding sequ
	26	13.2	66.0	60	24	ABN37006	Human spliced tran
	27	13.2	66.0	60	24	ABN49957	Human spliced tran
c	28	13.2	66.0	65	24	ABN29622	Rat spliced transc
	29	13.2	66.0	65	24	ABN56709	Mouse spliced tran
c	30	13.2	66.0	75	21	AAC30437	Human secreted pro
	31	13.2	66.0	76	24	ABK12147	Rabbit tRNAlys3 se
	32	13.2	66.0	97	21	AAC13733	Human secreted pro
c	33	12.8	64.0	21	20	AAZ26930	Human chromosome 1
c	34	12.8	64.0	40	21	AAZ89515	Human LAM-PCR prim
	35	12.8	64.0	60	21	AAA70391	Site directed muta
	36	12.8	64.0	60	24	ABL36114	E. coli connectron
c	37	12.8	64.0	65	24	ABN53343	Mouse spliced tran
	38	12.8	64.0	71	24	ABL36106	E. coli connectron
	39	12.8	64.0	71	24	ABL36111	E. coli connectron
	40	12.8	64.0	86	24	ABL36109	E. coli connectron
	41	12.8	64.0	86	24	ABL36113	E. coli connectron
c	42	12.8	64.0	88	21	AAC32095	Human secreted pro
	43	12.8	64.0	98	24	ABL36108	E. coli connectron
	44	12.8	64.0	98	24	ABL36110	E. coli connectron
	45	12.6	63.0	25	22	AAH74723	PCR primer used to

ALIGNMENTS

RESULT 1
AAZ31436
ID AAZ31436 standard; DNA; 20 BP.
XX
AC AAZ31436;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3606.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 16; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 3 A; 3 C; 11 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTAGTTCAGGCGGGAGCG 20
||||||||||||||||||||
Db 1 GGGTAGTTCAGGCGGGAGCG 20

RESULT 2
AAF83316/c
ID AAF83316 standard; DNA; 30 BP.
XX
AC AAF83316;
XX
DT 09-JUL-2001 (first entry)
XX
DE Human neuropilin-1 (NP-1) DNA amplifying sense primer BJL-259.
XX
KW Receptor protein; vascular endothelial growth factor receptor-2; NP-1;
KW VEGFR-2; neuropilin-1; co-receptor; human; angiogenic; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200131346-A2.
XX
PD 03-MAY-2001.
XX
PF 26-OCT-2000; 2000WO-US29579.
XX
PR 28-OCT-1999; 99US-0162367.
XX
PA (PROC) PROCTER & GAMBLE CO.
XX
PI Rosenbaum JS, Whitaker GB, Limberg BJ;
XX
DR WPI; 2001-308686/32.
XX
PT Determining compounds which bind to a complex comprising vascular
PT endothelial growth factor receptor-2 and Neuropilin-1 to provide
PT superior pro- and anti-angiogenic agents -
XX
PS Example 1; Fig 3; 82pp; English.
XX
CC The invention relates to determining whether a compound is capable of
CC binding to a receptor protein complex comprising a vascular endothelial
CC growth factor receptor-2 (VEGFR-2) receptor protein and a neuropilin-1
CC (NP-1) receptor protein. One method comprises introducing a sample
CC comprising the compound to the receptor protein and allowing the
CC compound to bind to the complex. Signaling through VEGFR-2 is enhanced
CC in the presence of the NP-1 co-receptor. The methods of the invention
CC can be used for identifying novel pro- and anti-angiogenic compounds.
CC Sequences AAF83314-16 represents PCR primers used for isolating human
CC NP-1 DNA.
XX
SQ Sequence 30 BP; 5 A; 11 C; 8 G; 6 T; 0 other;

Query Match 90.0%; Score 18; DB 22; Length 30;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GTAGTTCAGGCGGGAGCG 20
||||||||||||||||||
Db 30 GTAGTTCAGGCGGGAGCG 13

RESULT 3
AAI74524/c
ID AAI74524 standard; DNA; 51 BP.
XX
AC AAI74524;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ:1465.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 502; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 51 BP; 10 A; 19 C; 10 G; 12 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 GGGTAGTTCAGGCGGGAGC 19
           |||| || ||| |||||||
Db      28 GGGTGGTACAGCCGGGAGC 10
```

RESULT 4
ABN52875/c
ID ABN52875 standard; DNA; 65 BP.
XX
AC ABN52875;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:25623.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 25623; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 19 A; 19 C; 13 G; 14 T; 0 other;

Query Match 71.0%; Score 14.2; DB 24; Length 65;
Best Local Similarity 84.2%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 GGTAGTTCAGGCGGGAGCG 20
           ||||||||  | |||||||
Db      65 GGTAGTTCCTGTGGGAGCG 47
```

RESULT 5
AAF75497/c
ID AAF75497 standard; DNA; 95 BP.
XX
AC AAF75497;
XX
DT 14-MAY-2001 (first entry)
XX
DE Codon-optimised HPV6 E2 fragment 6M.
XX
KW Human papillomavirus; HPV; HPV16; HPV6a; HPV18; L1; E2; E7; E1;
KW antiviral; immunostimulant; vaccine; immunogen; infection; ss.
XX
OS Human papillomavirus.
OS Synthetic.
XX
PN WO200114416-A2.
XX
PD 01-MAR-2001.
XX
PF 21-AUG-2000; 2000WO-US22932.
XX
PR 25-AUG-1999; 99US-0150728.
PR 07-JUN-2000; 2000US-0210143.
XX
PA (MERI) MERCK & CO INC.
XX
PI Neeper MP, McClements WL, Jansen KU, Schultz LD, Chen L, Wang X;
XX
DR WPI; 2001-218428/22.
XX
PT Novel synthetic polynucleotide encoding human papillomavirus (HPV)
PT protein or mutated HPV protein useful as anti-HPV vaccines, comprises
PT optimized-codons for expression of the viral proteins in human host
PT cells -
XX
PS Example 6; Fig 23; 119pp; English.
XX
CC The present sequence is an oligomer which was used in the assembly of
CC one of a number of synthetic polynucleotides that encode a human
CC papillomavirus (HPV) protein, or a mutated form of a HPV protein. The
CC mutated HPV proteins have reduced protein function as compared to wild
CC type proteins but maintain immunogenicity. The proteins comprise codons
CC for optimised expression in humans. The polynucleotides are useful as a
CC vaccine which provides effective immunoprophylaxis against
CC papillomavirus infection through stimulation of neutralising antibody
CC and cell-mediated immunity.
XX
SQ Sequence 95 BP; 23 A; 36 C; 23 G; 13 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 95;
Best Local Similarity 84.2%; Pred. No. 1.2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 GGTAGTTCAGGCGGGAGCG 20
           | | ||||||||| ||||
Db      92 GATCGTTCAGGCGGTAGCG 74
```

RESULT 6
AAF75498
ID AAF75498 standard; DNA; 96 BP.
XX
AC AAF75498;
XX
DT 14-MAY-2001 (first entry)
XX

DE Codon-optimised HPV6 E2 fragment 6N.
XX
KW Human papillomavirus; HPV; HPV16; HPV6a; HPV18; L1; E2; E7; E1;
KW antiviral; immunostimulant; vaccine; immunogen; infection; ss.
XX
OS Human papillomavirus.
OS Synthetic.
XX
PN WO200114416-A2.
XX
PD 01-MAR-2001.
XX
PF 21-AUG-2000; 2000WO-US22932.
XX
PR 25-AUG-1999; 99US-0150728.
PR 07-JUN-2000; 2000US-0210143.
XX
PA (MERI) MERCK & CO INC.
XX
PI Neeper MP, McClements WL, Jansen KU, Schultz LD, Chen L, Wang X;
XX
DR WPI; 2001-218428/22.
XX
PT Novel synthetic polynucleotide encoding human papillomavirus (HPV)
PT protein or mutated HPV protein useful as anti-HPV vaccines, comprises
PT optimized-codons for expression of the viral proteins in human host
PT cells -
XX
PS Example 6; Fig 23; 119pp; English.
XX
CC The present sequence is an oligomer which was used in the assembly of
CC one of a number of synthetic polynucleotides that encode a human
CC papillomavirus (HPV) protein, or a mutated form of a HPV protein. The
CC mutated HPV proteins have reduced protein function as compared to wild
CC type proteins but maintain immunogenicity. The proteins comprise codons
CC for optimised expression in humans. The polynucleotides are useful as a
CC vaccine which provides effective immunoprophylaxis against
CC papillomavirus infection through stimulation of neutralising antibody
CC and cell-mediated immunity.
XX
SQ Sequence 96 BP; 13 A; 23 C; 40 G; 20 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 96;
Best Local Similarity 84.2%; Pred. No. 1.2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GGTAGTTCAGGCGGGAGCG 20
| | |||||||||| ||||
Db 73 GATCGTTCAGGCGGTAGCG 91

RESULT 7
ABK53798
ID ABK53798 standard; DNA; 30 BP.
XX
AC ABK53798;
XX
DT 05-JUN-2002 (first entry)
XX
DE DMS:acceptor oxidoreductase, PCR primer #44.
XX
KW DMS:acceptor oxidoreductase; dimethyl sulphide; sulphoxide;
KW prochiral organic sulphide; sulphoxide enantiomer; primer;
KW chiral drug production; optically-active functional drug; ss.
XX
OS Rhodovulum sulfidophilum.
XX
PN WO200216570-A1.
XX
PD 28-FEB-2002.
XX
PF 21-AUG-2001; 2001WO-AU01033.
XX
PR 21-AUG-2000; 2000AU-0009559.
XX
PA (UYQU) UNIV QUEENSLAND.
XX
PI McDevitt CA, McEwan AG;
XX
DR WPI; 2002-280922/32.
XX
PT New recombinant dimethyl sulphide:acceptor oxidoreductase or its
PT subunits, useful for oxidising prochiral organic sulphides to form
PT sulphoxide enantiomers for chiral drug synthesis -
XX
PS Claim 15; Page 47; 66pp; English.
XX
CC The invention relates to a recombinant dimethyl sulphide (DMS):acceptor
CC oxidoreductase (I) or its subunit selected from recombinant alpha,
CC beta, delta and gamma subunits. (I) is useful for oxidising prochiral
CC organic sulphides to form sulphoxide enantiomers for chiral drug
CC synthesis. (I) is expressed in a transformed bacterium. The
CC enantiomer formed is useful for producing a chiral drug. (I) is useful
CC for synthesis of optically-active functional groups of drug. DNA encoding
CC (I) is useful for producing a strain of DMS:acceptor oxidoreductase-
CC deficient Rhodovulum sulfidophilum, which is useful in whole-cell
CC reaction, where DMS:acceptor oxidoreductase activity is unwanted.
CC ABK53751-ABK53805 represent R. sulfidophilum DMS:acceptor oxidoreductase
CC subunit coding sequences and PCR primers of the invention.
XX
SQ Sequence 30 BP; 3 A; 9 C; 12 G; 6 T; 0 other;

Query Match 68.0%; Score 13.6; DB 24; Length 30;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GGGTAGTTCAGGCGGGAGCG 20
||||| ||||||||| ||
Db 6 GGGTACCTCAGGCGGGTTCG 25

RESULT 8
AAZ09153
ID AAZ09153 standard; DNA; 63 BP.
XX
AC AAZ09153;
XX
DT 18-OCT-1999 (first entry)
XX
DE Hairpin type nucleic acid probe 4.
XX
KW Hairpin; probe; separation; autocomplementary stem loop; marker; ss.
XX
OS Synthetic.
XX
PN JP11206381-A.
XX
PD 03-AUG-1999.
XX
PF 29-JAN-1998; 98JP-0016987.
XX
PR 29-JAN-1998; 98JP-0016987.
XX
PA (AISH-) AISHIN COSMOS KENKYUSHO KK.
XX
DR WPI; 1999-486357/41.
XX
PT Use of hairpin type nucleic acid probe - useful for detecting and
PT isolating nucleic acid in short time free from denaturation or
PT mutation
XX
PS Disclosure; Page 13; 15pp; Japanese.
XX
CC This invention describes the use of a hairpin type nucleic acid probe
CC molecule for separation of a target nucleic acid having a desired
CC nucleotide sequence composed of a 1st nucleotide part having an

```
CC   autocomplementary stem loop structure, a 2nd nucleotide part, especially
CC   composed of at least 15 bases, having a single stranded structure
CC   extended from one terminal of the 1st nucleotide part and a marker for
CC   the separation, especially biotin or digoxigenin, being contained in the
CC   1st and/or 2nd part(s) and at least partially complementary to the above
CC   mentioned nucleotide sequences and further the marker for separation is
CC   capable of specific binding with a particular substance. Detection and
CC   isolation of an intact target nucleic acid is available in a short
CC   period of time free from denaturation or mutation, without requiring
CC   construction of DNA library or its screening. AAZ09150-Z09155 are DNA
CC   probes used in the method of the invention.
XX
SQ   Sequence 63 BP; 4 A; 20 C; 25 G; 14 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 20;  Length 63;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAGCG 20
           ||| || |||||||  |||
Db      30 GGGCAGGTCAGGCGGCCGCG 49


RESULT 9
ABK87514
ID   ABK87514 standard; RNA; 76 BP.
XX
AC   ABK87514;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Human immunodeficiency virus (HIV) transfer RNA (tRNA)Lys1,2.
XX
KW   Transfer RNA; tRNA; transfer ribonucleic acid; retrovirus;
KW   reverse transcriptase priming; human immunodeficiency virus; HIV;
KW   simian immunodeficiency virus; tRNALys3; aminoacyl tRNA synthetase;
KW   LysRS; tRNALys1,2; ss.
XX
OS   Human immunodeficiency virus type 1.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       10..25
FT                   /*tag= a
FT   modified_base   10
FT                   /*tag= b
FT                   /mod_base= G
FT                   /note= "2-methyl Guanine"
FT   stem_loop       27..43
FT                   /*tag= c
FT   modified_base   46
FT                   /*tag= d
FT                   /mod_base= G
FT                   /note= "7-methyl Guanine"
FT   modified_base   48
FT                   /*tag= e
FT                   /mod_base= C
FT                   /note= "5-methyl Cytosine"
FT   modified_base   49
FT                   /*tag= f
FT                   /mod_base= C
FT                   /note= "5-methyl Cytosine"
FT   modified_base   58
FT                   /*tag= g
FT                   /mod_base= A
FT                   /note= "1-methyl Adenine"
FT   stem_loop       49..65
FT                   /*tag= h
XX
PN   WO200244349-A2.
XX
PD   06-JUN-2002.
XX
PF   27-NOV-2001; 2001WO-CA01700.
XX
PR   28-NOV-2000; 2000US-253146P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Cen S,  Kleiman L;
XX
DR   WPI; 2002-508511/54.
XX
PT   Modulating incorporation of tRNA involved in reverse transcriptase
PT   priming into retrovirus e.g. human immunodeficiency virus tRNA-Lys3
PT   comprises modulating activity and/or level of cognate aminoacyl tRNA
PT   synthetase -
XX
PS   Disclosure; Fig 6; 121pp; English.
XX
CC   The invention describes a method of modulating incorporation of transfer
CC   ribonucleic acid (tRNA) involved in reverse transcriptase priming into
CC   retroviral virion such as human immunodeficiency virus or simian
CC   immunodeficiency virus, e.g. tRNA(Lys3). The method comprises modulating
CC   the activity and/or level of cognate aminoacyl tRNA synthetase e.g.
CC   LysRS, where the level positively correlates with an incorporation of the
CC   tRNA into the virion. The invention also describes a method useful for
CC   targeting a molecule into HIV by providing the molecule linked to a
CC   sufficient number of aminoacyl tRNA synthetases, preferably LysRS. The
CC   invention also describes a method useful for screening and selecting an
CC   agent that modulates the incorporation of tRNA (tRNA(Lys), preferably
CC   tRNA(Lys3)) and/or cognate aminoacyl tRNA synthetase (preferably Lys)
CC   into a retroviral virion such as HIV or related viruses. The invention
CC   describes the use of a molecule for interfering with incorporation of
CC   native tRNA (preferably tRNA(Lys), more preferably tRNA(Lys3)) involved
CC   in reverse transcriptase (RT) priming and/or its native cognate aminoacyl
CC   tRNA synthetase (preferably LysRS) into a retroviral virion such as HIV
CC   or related viruses, preferably HIV-1. This sequence represents the Human
CC   immunodeficiency virus type 1 (HIV-1) isoacceptor tRNA Lys1,2, the
CC   function of which is not yet known.
XX
SQ   Sequence 76 BP; 11 A; 23 C; 26 G; 7 U; 9 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 76;
  Best Local Similarity   65.0%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAGCG 20
           || :|| :|||:|||:|| |
Db       5 GGCUAGCUCAGDCGGDAGAG 24


RESULT 10
ABK87515
ID   ABK87515 standard; RNA; 76 BP.
XX
AC   ABK87515;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Human immunodeficiency virus (HIV) transfer RNA (tRNA)Lys3.
XX
KW   Transfer RNA; tRNA; transfer ribonucleic acid; retrovirus;
KW   reverse transcriptase priming; human immunodeficiency virus; HIV;
KW   simian immunodeficiency virus; tRNALys3; aminoacyl tRNA synthetase;
KW   LysRS; tRNALys3; ss.
XX
OS   Human immunodeficiency virus type 1.
XX
FH   Key             Location/Qualifiers
FT   stem_loop       10..25
FT                   /*tag= a
FT   modified_base   10
FT                   /*tag= b
FT                   /mod_base= G
FT                   /note= "2-methyl Guanine"
FT   stem_loop       27..43
```

```
FT                   /*tag=  c
FT   modified_base   46
FT                   /*tag=  d
FT                   /mod_base= G
FT                   /note= "7-methyl Guanine"
FT   modified_base   48
FT                   /*tag=  e
FT                   /mod_base= C
FT                   /note= "5-methyl Cytosine"
FT   modified_base   49
FT                   /*tag=  f
FT                   /mod_base= C
FT                   /note= "5-methyl Cytosine"
FT   modified_base   58
FT                   /*tag=  g
FT                   /mod_base= A
FT                   /note= "1-methyl Adenine"
FT   stem_loop       49..65
FT                   /*tag=  h
XX
PN   WO200244349-A2.
XX
PD   06-JUN-2002.
XX
PF   27-NOV-2001; 2001WO-CA01700.
XX
PR   28-NOV-2000; 2000US-253146P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Cen S,  Kleiman L;
XX
DR   WPI; 2002-508511/54.
XX
PT   Modulating incorporation of tRNA involved in reverse transcriptase
PT   priming into retrovirus e.g. human immunodeficiency virus tRNA-Lys3
PT   comprises modulating activity and/or level of cognate aminoacyl tRNA
PT   synthetase  -
XX
PS   Disclosure; Fig 6; 121pp; English.
XX
CC   The invention describes a method of modulating incorporation of transfer
CC   ribonucleic acid (tRNA) involved in reverse transcriptase priming into
CC   retroviral virion such as human immunodeficiency virus or simian
CC   immunodeficiency virus, e.g. tRNA(Lys3). The method comprises modulating
CC   the activity and/or level of cognate aminoacyl tRNA synthetase e.g.
CC   LysRS, where the level positively correlates with an incorporation of the
CC   tRNA into the virion. The invention also describes a method useful for
CC   targeting a molecule into HIV by providing the molecule linked to a
CC   sufficient number of aminoacyl tRNA synthetases, preferably LysRS. The
CC   invention also describes a method useful for screening and selecting an
CC   agent that modulates the incorporation of tRNA (tRNA(Lys), preferably
CC   tRNA(Lys3)) and/or cognate aminoacyl tRNA synthetase (preferably Lys)
CC   into a retroviral virion such as HIV or related viruses. The invention
CC   describes the use of a molecule for interfering with incorporation of
CC   native tRNA (preferably tRNA(Lys), more preferably tRNA(Lys3)) involved
CC   in reverse transcriptase (RT) priming and/or its native cognate aminoacyl
CC   tRNA synthetase (preferably LysRS) into a retroviral virion such as HIV
CC   or related viruses, preferably HIV-1. This sequence represents the Human
CC   immunodeficiency virus type 1 (HIV-1) isoacceptor tRNA Lys3, used as the
CC   primer for reverse transcriptase catalysed synthesis of minus strand DNA.
XX
SQ   Sequence 76 BP; 14 A; 20 C; 23 G; 10 U; 9 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 76;
  Best Local Similarity   65.0%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAGCG 20
           || :|| :|||:|||:|| |
Db       5 GGAUAGCUCAGDCGGDAGAG 24
```

```
RESULT 11
ABK87516
ID   ABK87516 standard; RNA; 76 BP.
XX
AC   ABK87516;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Human immunodeficiency virus (HIV) transfer RNA 35(tRNA)Lys3.
XX
KW   Transfer RNA; tRNA; transfer ribonucleic acid; retrovirus;
KW   reverse transcriptase priming; human immunodeficiency virus; HIV;
KW   simian immunodeficiency virus; tRNALys3; aminoacyl tRNA synthetase;
KW   LysRS; tRNALys3; mutant; ss.
XX
OS   Human immunodeficiency virus type 1.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   mutation        replace(35,U)
FT                   /*tag=  a
XX
PN   WO200244349-A2.
XX
PD   06-JUN-2002.
XX
PF   27-NOV-2001; 2001WO-CA01700.
XX
PR   28-NOV-2000; 2000US-253146P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Cen S,  Kleiman L;
XX
DR   WPI; 2002-508511/54.
XX
PT   Modulating incorporation of tRNA involved in reverse transcriptase
PT   priming into retrovirus e.g. human immunodeficiency virus tRNA-Lys3
PT   comprises modulating activity and/or level of cognate aminoacyl tRNA
PT   synthetase  -
XX
PS   Example 4; Page -; 121pp; English.
XX
CC   The invention describes a method of modulating incorporation of transfer
CC   ribonucleic acid (tRNA) involved in reverse transcriptase priming into
CC   retroviral virion such as human immunodeficiency virus or simian
CC   immunodeficiency virus, e.g. tRNA(Lys3). The method comprises modulating
CC   the activity and/or level of cognate aminoacyl tRNA synthetase e.g.
CC   LysRS, where the level positively correlates with an incorporation of the
CC   tRNA into the virion. The invention also describes a method useful for
CC   targeting a molecule into HIV by providing the molecule linked to a
CC   sufficient number of aminoacyl tRNA synthetases, preferably LysRS. The
CC   invention also describes a method useful for screening and selecting an
CC   agent that modulates the incorporation of tRNA (tRNA(Lys), preferably
CC   tRNA(Lys3)) and/or cognate aminoacyl tRNA synthetase (preferably Lys)
CC   into a retroviral virion such as HIV or related viruses. The invention
CC   describes the use of a molecule for interfering with incorporation of
CC   native tRNA (preferably tRNA(Lys), more preferably tRNA(Lys3)) involved
CC   in reverse transcriptase (RT) priming and/or its native cognate aminoacyl
CC   tRNA synthetase (preferably LysRS) into a retroviral virion such as HIV
CC   or related viruses, preferably HIV-1. This sequence represents a mutant
CC   of the Human immunodeficiency virus type 1 (HIV-1) isoacceptor tRNA Lys3,
CC   used as the primer for reverse transcriptase catalysed synthesis of minus
CC   strand DNA.
CC   Note: This sequence does not appear in the specification but has been
CC   created using the wild type sequence shown in ABK87515 using information
CC   given in the specification.
XX
SQ   Sequence 76 BP; 14 A; 20 C; 24 G; 9 U; 9 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 76;
  Best Local Similarity   65.0%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy       1 GGGTAGTTCAGGCGGGAGCG 20
           || :|| :|||:|||:|| |
Db       5 GGAUAGCUCAGDCGGDAGAG 24


RESULT 12
ABK87517
ID   ABK87517 standard; RNA; 76 BP.
XX
AC   ABK87517;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Human immunodeficiency virus (HIV) transfer RNA 34, 35(tRNA)Lys3.
XX
KW   Transfer RNA; tRNA; transfer ribonucleic acid; retrovirus;
KW   reverse transcriptase priming; human immunodeficiency virus; HIV;
KW   simian immunodeficiency virus; tRNALys3; aminoacyl tRNA synthetase;
KW   LysRS; tRNALys3; mutant; ss.
XX
OS   Human immunodeficiency virus type 1.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   mutation        replace(34,S)
FT                   /*tag=  a
FT   mutation        replace(35,U)
FT                   /*tag=  b
XX
PN   WO200244349-A2.
XX
PD   06-JUN-2002.
XX
PF   27-NOV-2001; 2001WO-CA01700.
XX
PR   28-NOV-2000; 2000US-253146P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Cen S,  Kleiman L;
XX
DR   WPI; 2002-508511/54.
XX
PT   Modulating incorporation of tRNA involved in reverse transcriptase
PT   priming into retrovirus e.g. human immunodeficiency virus tRNA-Lys3
PT   comprises modulating activity and/or level of cognate aminoacyl tRNA
PT   synthetase  -
XX
PS   Example 4; Page -; 121pp; English.
XX
CC   The invention describes a method of modulating incorporation of transfer
CC   ribonucleic acid (tRNA) involved in reverse transcriptase priming into
CC   retroviral virion such as human immunodeficiency virus or simian
CC   immunodeficiency virus, e.g. tRNA(Lys3). The method comprises modulating
CC   the activity and/or level of cognate aminoacyl tRNA synthetase e.g.
CC   LysRS, where the level positively correlates with an incorporation of the
CC   tRNA into the virion. The invention also describes a method useful for
CC   targeting a molecule into HIV by providing the molecule linked to a
CC   sufficient number of aminoacyl tRNA synthetases, preferably LysRS. The
CC   invention also describes a method useful for screening and selecting an
CC   agent that modulates the incorporation of tRNA (tRNA(Lys), preferably
CC   tRNA(Lys3)) and/or cognate aminoacyl tRNA synthetase (preferably Lys)
CC   into a retroviral virion such as HIV or related viruses. The invention
CC   describes the use of a molecule for interfering with incorporation of
CC   native tRNA (preferably tRNA(Lys), more preferably tRNA(Lys3)) involved
CC   in reverse transcriptase (RT) priming and/or its native cognate aminoacyl
CC   tRNA synthetase (preferably LysRS) into a retroviral virion such as HIV
CC   or related viruses, preferably HIV-1. This sequence represents a mutant
CC   of the Human immunodeficiency virus type 1 (HIV-1) isoacceptor tRNA Lys3,
CC   used as the primer for reverse transcriptase catalysed synthesis of minus
CC   strand DNA.
CC   Note: This sequence does not appear in the specification but has been
CC   created using the wild type sequence shown in ABK87515 using information
CC   given in the specification.
XX
SQ   Sequence 76 BP; 14 A; 21 C; 24 G; 9 U; 8 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 76;
  Best Local Similarity   65.0%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAGCG 20
           || :|| :|||:|||:|| |
Db       5 GGAUAGCUCAGDCGGDAGAG 24


RESULT 13
ABK87518
ID   ABK87518 standard; RNA; 76 BP.
XX
AC   ABK87518;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Human immunodeficiency virus (HIV) transfer RNA 35, 36(tRNA)Lys3.
XX
KW   Transfer RNA; tRNA; transfer ribonucleic acid; retrovirus;
KW   reverse transcriptase priming; human immunodeficiency virus; HIV;
KW   simian immunodeficiency virus; tRNALys3; aminoacyl tRNA synthetase;
KW   LysRS; tRNALys3; mutant; ss.
XX
OS   Human immunodeficiency virus type 1.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   mutation        replace(35,U)
FT                   /*tag=  a
FT   mutation        replace(36,U)
FT                   /*tag=  b
XX
PN   WO200244349-A2.
XX
PD   06-JUN-2002.
XX
PF   27-NOV-2001; 2001WO-CA01700.
XX
PR   28-NOV-2000; 2000US-253146P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Cen S,  Kleiman L;
XX
DR   WPI; 2002-508511/54.
XX
PT   Modulating incorporation of tRNA involved in reverse transcriptase
PT   priming into retrovirus e.g. human immunodeficiency virus tRNA-Lys3
PT   comprises modulating activity and/or level of cognate aminoacyl tRNA
PT   synthetase  -
XX
PS   Example 4; Page -; 121pp; English.
XX
CC   The invention describes a method of modulating incorporation of transfer
CC   ribonucleic acid (tRNA) involved in reverse transcriptase priming into
CC   retroviral virion such as human immunodeficiency virus or simian
CC   immunodeficiency virus, e.g. tRNA(Lys3). The method comprises modulating
CC   the activity and/or level of cognate aminoacyl tRNA synthetase e.g.
CC   LysRS, where the level positively correlates with an incorporation of the
CC   tRNA into the virion. The invention also describes a method useful for
CC   targeting a molecule into HIV by providing the molecule linked to a
CC   sufficient number of aminoacyl tRNA synthetases, preferably LysRS. The
CC   invention also describes a method useful for screening and selecting an
CC   agent that modulates the incorporation of tRNA (tRNA(Lys), preferably
CC   tRNA(Lys3)) and/or cognate aminoacyl tRNA synthetase (preferably Lys)
CC   into a retroviral virion such as HIV or related viruses. The invention
CC   describes the use of a molecule for interfering with incorporation of
```

```
CC   native tRNA (preferably tRNA(Lys), more preferably tRNA(Lys3)) involved
CC   in reverse transcriptase (RT) priming and/or its native cognate aminoacyl
CC   tRNA synthetase (preferably LysRS) into a retroviral virion such as HIV
CC   or related viruses, preferably HIV-1. This sequence represents a mutant
CC   of the Human immunodeficiency virus type 1 (HIV-1) isoacceptor tRNA Lys3,
CC   used as the primer for reverse transcriptase catalysed synthesis of minus
CC   strand DNA.
CC   Note: This sequence does not appear in the specification but has been
CC   created using the wild type sequence shown in ABK87515 using information
CC   given in the specification.
XX
SQ   Sequence 76 BP; 15 A; 20 C; 24 G; 8 U; 9 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 76;
  Best Local Similarity   65.0%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAGCG 20
           || :|| :|||:|||:|| |
Db       5 GGAUAGCUCAGDCGGDAGAG 24


RESULT 14
ABK87519
ID   ABK87519 standard; RNA; 76 BP.
XX
AC   ABK87519;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Human immunodeficiency virus (HIV) transfer RNA 34, 35, 36(tRNA)Lys3.
XX
KW   Transfer RNA; tRNA; transfer ribonucleic acid; retrovirus;
KW   reverse transcriptase priming; human immunodeficiency virus; HIV;
KW   simian immunodeficiency virus; tRNALys3; aminoacyl tRNA synthetase;
KW   LysRS; tRNALys3; mutant; ss.
XX
OS   Human immunodeficiency virus type 1.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   mutation        replace(34,S)
FT                   /*tag=  a
FT   mutation        replace(35,U)
FT                   /*tag=  a
FT   mutation        replace(36,U)
FT                   /*tag=  b
XX
PN   WO200244349-A2.
XX
PD   06-JUN-2002.
XX
PF   27-NOV-2001; 2001WO-CA01700.
XX
PR   28-NOV-2000; 2000US-253146P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Cen S,  Kleiman L;
XX
DR   WPI; 2002-508511/54.
XX
PT   Modulating incorporation of tRNA involved in reverse transcriptase
PT   priming into retrovirus e.g. human immunodeficiency virus tRNA-Lys3
PT   comprises modulating activity and/or level of cognate aminoacyl tRNA
PT   synthetase  -
XX
PS   Example 4; Page -; 121pp; English.
XX
CC   The invention describes a method of modulating incorporation of transfer
CC   ribonucleic acid (tRNA) involved in reverse transcriptase priming into
CC   retroviral virion such as human immunodeficiency virus or simian
CC   immunodeficiency virus, e.g. tRNA(Lys3). The method comprises modulating
CC   the activity and/or level of cognate aminoacyl tRNA synthetase e.g.
CC   LysRS, where the level positively correlates with an incorporation of the
CC   tRNA into the virion. The invention also describes a method useful for
CC   targeting a molecule into HIV by providing the molecule linked to a
CC   sufficient number of aminoacyl tRNA synthetases, preferably LysRS. The
CC   invention also describes a method useful for screening and selecting an
CC   agent that modulates the incorporation of tRNA (tRNA(Lys), preferably
CC   tRNA(Lys3)) and/or cognate aminoacyl tRNA synthetase (preferably Lys)
CC   into a retroviral virion such as HIV or related viruses. The invention
CC   describes the use of a molecule for interfering with incorporation of
CC   native tRNA (preferably tRNA(Lys), more preferably tRNA(Lys3)) involved
CC   in reverse transcriptase (RT) priming and/or its native cognate aminoacyl
CC   tRNA synthetase (preferably LysRS) into a retroviral virion such as HIV
CC   or related viruses, preferably HIV-1. This sequence represents a mutant
CC   of the Human immunodeficiency virus type 1 (HIV-1) isoacceptor tRNA Lys3,
CC   used as the primer for reverse transcriptase catalysed synthesis of minus
CC   strand DNA.
CC   Note: This sequence does not appear in the specification but has been
CC   created using the wild type sequence shown in ABK87515 using information
CC   given in the specification.
XX
SQ   Sequence 76 BP; 15 A; 21 C; 24 G; 8 U; 8 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 76;
  Best Local Similarity   65.0%;  Pred. No. 2.3e+03;
  Matches   13;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAGCG 20
           || :|| :|||:|||:|| |
Db       5 GGAUAGCUCAGDCGGDAGAG 24


RESULT 15
AAC09858/c
ID   AAC09858 standard; cDNA; 88 BP.
XX
AC   AAC09858;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 13933.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 13933; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
```

CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 88 BP; 11 A; 33 C; 24 G; 20 T; 0 other;

Query Match 68.0%; Score 13.6; DB 21; Length 88;
Best Local Similarity 80.0%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 GGGTAGTTCAGGCGGGAGCG 20
           ||| ||||||| ||||| |
Db      85 GGGAAGTTCAGAGGGGAGGG 66
```

Search completed: November 23, 2002, 06:28:44
Job time : 101.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-6
Perfect score: 20
Sequence: 1 gggtagttcaggcgggagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13.4	67.0	42	3	US-08-879-565-9	Sequence 9, Appli
2	13.4	67.0	63	1	US-08-530-492-6	Sequence 6, Appli
3	13.4	67.0	63	4	US-08-906-517-6	Sequence 6, Appli
4	13.4	67.0	73	1	US-07-982-712-12	Sequence 12, Appl
5	13.2	66.0	28	1	US-08-408-656-7	Sequence 7, Appli
6	13.2	66.0	28	5	PCT-US95-03602-7	Sequence 7, Appli
c 7	13.2	66.0	33	2	US-08-432-871C-14	Sequence 14, Appl
c 8	13.2	66.0	33	4	US-09-270-956-14	Sequence 14, Appl
c 9	12.8	64.0	21	4	US-09-262-773-105	Sequence 105, App
c 10	12.8	64.0	48	1	US-08-391-000-11	Sequence 11, Appl
c 11	12.8	64.0	48	2	US-08-741-931-11	Sequence 11, Appl
c 12	12.4	62.0	24	2	US-08-508-088-5	Sequence 5, Appli
c 13	12.4	62.0	24	2	US-09-009-925-5	Sequence 5, Appli
c 14	12.4	62.0	27	4	US-09-314-847A-13	Sequence 13, Appl
15	12.4	62.0	65	1	US-08-453-104-11	Sequence 11, Appl
16	12.4	62.0	65	2	US-08-694-824-11	Sequence 11, Appl
c 17	12.4	62.0	73	1	US-08-453-104-20	Sequence 20, Appl
c 18	12.4	62.0	73	2	US-08-694-824-20	Sequence 20, Appl
c 19	12.2	61.0	20	4	US-09-344-529-15	Sequence 15, Appl
20	12.2	61.0	20	4	US-09-907-843-30	Sequence 30, Appl
c 21	12.2	61.0	25	2	US-08-412-376-3	Sequence 3, Appli
c 22	12.2	61.0	25	2	US-08-859-998-1021	Sequence 1021, Ap
c 23	12.2	61.0	25	3	US-08-646-273-32	Sequence 32, Appl
c 24	12.2	61.0	25	4	US-09-225-928-1021	Sequence 1021, Ap
c 25	12.2	61.0	31	5	PCT-US92-10792-29	Sequence 29, Appl
c 26	12	60.0	21	1	US-08-009-263C-84	Sequence 84, Appl
c 27	12	60.0	21	3	US-08-838-715B-84	Sequence 84, Appl
c 28	12	60.0	31	1	US-08-706-037-2	Sequence 2, Appli
c 29	12	60.0	31	2	US-09-005-397-2	Sequence 2, Appli
30	12	60.0	44	4	US-08-918-288-66	Sequence 66, Appl
31	12	60.0	44	4	US-09-282-357-66	Sequence 66, Appl
c 32	12	60.0	67	2	US-08-400-115-19	Sequence 19, Appl
33	12	60.0	71	2	US-08-400-115-18	Sequence 18, Appl
34	12	60.0	76	1	US-08-345-623-3	Sequence 3, Appli
35	12	60.0	76	4	US-08-866-116A-19	Sequence 19, Appl
36	12	60.0	76	4	US-08-866-116A-20	Sequence 20, Appl
37	12	60.0	84	2	US-08-400-115-20	Sequence 20, Appl
38	12	60.0	87	4	US-08-866-116A-3	Sequence 3, Appli
39	12	60.0	92	5	PCT-US91-05808-1	Sequence 1, Appli
40	12	60.0	92	6	5168053-2	Patent No. 5168053
c 41	11.8	59.0	21	4	US-09-345-217-12	Sequence 12, Appl
c 42	11.8	59.0	22	1	US-08-510-696-1	Sequence 1, Appli
c 43	11.8	59.0	22	3	US-09-035-220-24	Sequence 24, Appl
c 44	11.8	59.0	22	5	PCT-US96-12455-1	Sequence 1, Appli
c 45	11.8	59.0	27	3	US-09-035-220-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-879-565-9
; Sequence 9, Application US/08879565A
; Patent No. 6093573
; GENERAL INFORMATION:
; APPLICANT: Beamer, Lesa J.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Eisenberg, David
; TITLE OF INVENTION: THREE-DIMENSIONAL STRUCTURE OF
; TITLE OF INVENTION: BACTERICIDAL/PERMEABILITY-INCREASING PROTEIN
; FILE REFERENCE: 1103/11034US01
; CURRENT APPLICATION NUMBER: US/08/879,565A
; CURRENT FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; OTHER INFORMATION: Oligonucleotide complementary to SEQ ID NO:8
US-08-879-565-9

Query Match 67.0%; Score 13.4; DB 3; Length 42;
Best Local Similarity 93.3%; Pred. No. 4.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      6 GTTCAGGCGGGAGCG 20
          ||||||||||||| ||
Db      3 GTTCAGGCGGGACCG 17
```

RESULT 2
US-08-530-492-6
; Sequence 6, Application US/08530492
; Patent No. 5689052
; GENERAL INFORMATION:
; APPLICANT: Brown, Sherri M.
; APPLICANT: Dean, Duff A.
; APPLICANT: Fromm, Michael E.
; APPLICANT: Sanders, Patricia R.
; TITLE OF INVENTION: Synthetic DNA Sequences Having Enhanced
; TITLE OF INVENTION: Expression in Monocotyledonous Plants and Method For
; TITLE OF INVENTION: Preparation Thereof
; NUMBER OF SEQUENCES: 164
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Parkway No. 5689052th
; CITY: St. Louis
; STATE: Missouri

```
;       COUNTRY:  USA
;       ZIP:  63198
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/530,492
;       FILING DATE:
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/172,333
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hoerner Jr., Dennis R.
;       REGISTRATION NUMBER:  30,914
;       REFERENCE/DOCKET NUMBER:  38-21(10605)A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (314)537-6099
;       TELEFAX:  (314)537-6047
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  63 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (synthetic)
US-08-530-492-6

  Query Match             67.0%;  Score 13.4;  DB 1;  Length 63;
  Best Local Similarity   93.3%;  Pred. No. 4.3e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 AGTTCAGGCGGGAGC 19
           ||||||||||  ||||
Db      26 AGTTCAGGCGCGAGC 40


RESULT 3
US-08-906-517-6
; Sequence 6, Application US/08906517
; Patent No. 6180774
;  GENERAL INFORMATION:
;     APPLICANT:  Brown, Sherri M.
;     APPLICANT:  Dean, Duff A.
;     APPLICANT:  Fromm, Michael E.
;     APPLICANT:  Sanders, Patricia R.
;     TITLE OF INVENTION:  Synthetic DNA Sequences Having Enhanced
;     TITLE OF INVENTION:  Expression in Monocotyledonous Plants and Method For
;     TITLE OF INVENTION:  Preparation Thereof
;     NUMBER OF SEQUENCES:  164
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Arnold, White & Durkee
;       STREET:  P.O. Box 4433
;       CITY:  Houston
;       STATE:  TX
;       COUNTRY:  USA
;       ZIP:  77210-4433
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/906,517
;       FILING DATE:  Concurrently Herewith
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kitchell, Barbara S.
;       REGISTRATION NUMBER:  33,928
;       REFERENCE/DOCKET NUMBER:  MOBT:170
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512-418-3000
;       TELEFAX:  512-474-7577
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  63 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-906-517-6

  Query Match             67.0%;  Score 13.4;  DB 4;  Length 63;
  Best Local Similarity   93.3%;  Pred. No. 4.3e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 AGTTCAGGCGGGAGC 19
           ||||||||||  ||||
Db      26 AGTTCAGGCGCGAGC 40


RESULT 4
US-07-982-712-12
; Sequence 12, Application US/07982712
; Patent No. 5436391
;  GENERAL INFORMATION:
;     APPLICANT:  Hideya FUJIMOTO, Kimiko ITOH
;     APPLICANT:  Mikihiro YAMAMOTO, and Ko SHIMAMOTO
;     TITLE OF INVENTION:  Insecticidal Protein-encoding Gene, Gramineous
;     TITLE OF INVENTION:  Plants Transformed with the Gene, and Production Thereof
;     NUMBER OF SEQUENCES:  35
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Wenderoth, Lind & Ponack
;       STREET:  805 Fifteenth Street, N.W., #700
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20005
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 144 mb
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  Wordperfect 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/982,712
;       FILING DATE:  19921127
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Warren M. Cheek, Jr.
;       REGISTRATION NUMBER:  33,367
;       REFERENCE/DOCKET NUMBER:
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-371-8850
;       TELEFAX:
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  73 bases
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  other nucleic acid
US-07-982-712-12

  Query Match             67.0%;  Score 13.4;  DB 1;  Length 73;
  Best Local Similarity   93.3%;  Pred. No. 4.3e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 AGTTCAGGCGGGAGC 19
           ||||||||  ||||||
Db      29 AGTTCAGGAGGGAGC 43
```

```
RESULT 5
US-08-408-656-7
; Sequence 7, Application US/08408656
; Patent No. 5674683
;  GENERAL INFORMATION:
;    APPLICANT:  Kool, Eric T.
;    TITLE OF INVENTION:  STEM-LOOP AND CIRCULAR OLIGONUCLEOTIDES
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/408,656
;      FILING DATE:  21-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9373
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  28 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  circular
US-08-408-656-7

  Query Match             66.0%;  Score 13.2;  DB 1;  Length 28;
  Best Local Similarity   83.3%;  Pred. No. 5e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAG 18
           |||| || |||| |||||
Db       6 GGGTGGTCCAGGAGGGAG 23


RESULT 6
PCT-US95-03602-7
; Sequence 7, Application PC/TUS9503602
;  GENERAL INFORMATION:
;    APPLICANT:  Research Corporation Technologies, Inc.
;    TITLE OF INVENTION:  STEM-LOOP AND CIRCULAR OLIGONUCLEOTIDES
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03602
;      FILING DATE:  21-MAR-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9373
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  28 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  circular
PCT-US95-03602-7

  Query Match             66.0%;  Score 13.2;  DB 5;  Length 28;
  Best Local Similarity   83.3%;  Pred. No. 5e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAG 18
           |||| || |||| |||||
Db       6 GGGTGGTCCAGGAGGGAG 23


RESULT 7
US-08-432-871C-14/c
; Sequence 14, Application US/08432871C
; Patent No. 5877010
;  GENERAL INFORMATION:
;    APPLICANT:  Loeb, Lawrence A.
;    APPLICANT:  Black, Margaret E.
;    TITLE OF INVENTION:  THYMIDINE KINASE MUTANTS
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/432,871C
;      FILING DATE:  02-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  240052.409C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;      TELEX:  3723836
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  33 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..33
US-08-432-871C-14

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 33;
  Best Local Similarity   83.3%;  Pred. No. 5.1e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy       1 GGGTAGTTCAGGCGGGAG 18
           ||||||  |||| |||||
Db      26 GGGTAGCACAGGAGGGAG 9


RESULT 8
US-09-270-956-14/c
; Sequence 14, Application US/09270956
; Patent No. 6451571
;  GENERAL INFORMATION:
;    APPLICANT:  Loeb, Lawrence A.
;    APPLICANT:  Black, Margaret E.
;    TITLE OF INVENTION:  THYMIDINE KINASE MUTANTS
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/270,956
;      FILING DATE:  17-MAR-1999
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  240052.409C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;      TELEX:  3723836
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  33 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(1..33)
US-09-270-956-14

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 33;
  Best Local Similarity   83.3%;  Pred. No. 5.1e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAG 18
           ||||||  |||| |||||
Db      26 GGGTAGCACAGGAGGGAG 9


RESULT 9
US-09-262-773-105/c
; Sequence 105, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
;  APPLICANT: Ballinger, Dennis G.
;  APPLICANT: Ding, Wei
;  APPLICANT: Wagner, Susanne
;  APPLICANT: Hess, Mark A.
;  TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
;  FILE REFERENCE: Myriad 3
;  CURRENT APPLICATION NUMBER: US/09/262,773
;  CURRENT FILING DATE: 1999-03-04
;  NUMBER OF SEQ ID NOS: 210
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: primer
US-09-262-773-105

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 7.7e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GTAGTTCAGGCGGGAG 18
           ||||| ||||| ||||
Db      19 GTAGTCCAGGCAGGAG 4


RESULT 10
US-08-391-000-11/c
; Sequence 11, Application US/08391000
; Patent No. 5723752
;  GENERAL INFORMATION:
;    APPLICANT:  HOUTZ, Robert L.
;    TITLE OF INVENTION:  CLONING AND DEVELOPMENTAL EXPRESSION OF
;    TITLE OF INVENTION:  PEA RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE
;    TITLE OF INVENTION:  SUBUNIT N-METHYLTRANSFERASE
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/391,000
;      FILING DATE:  21-FEB-1995
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meuth, Donna M.
;      REGISTRATION NUMBER:  36,607
;      REFERENCE/DOCKET NUMBER:  028750-123
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-391-000-11

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 48;
  Best Local Similarity   87.5%;  Pred. No. 8.2e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TAGTTCAGGCGGGAGC 19
           |||||||| ||| |||
Db      17 TAGTTCAGCCGGTAGC 2


RESULT 11
US-08-741-931-11/c
; Sequence 11, Application US/08741931
; Patent No. 5866394
;  GENERAL INFORMATION:
;    APPLICANT:  HOUTZ, Robert L.
```

```
;    TITLE OF INVENTION:  CLONING AND DEVELOPMENTAL EXPRESSION OF
;    TITLE OF INVENTION:  PEA RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE
;    TITLE OF INVENTION:  SUBUNIT N-METHYLTRANSFERASE
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis
;      STREET:  P.O. Box 1404
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/741,931
;      FILING DATE:  31-OCT-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/391,000
;      FILING DATE:  21-FEB-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meuth, Donna M.
;      REGISTRATION NUMBER:  36,607
;      REFERENCE/DOCKET NUMBER:  028750-123
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-741-931-11

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 48;
  Best Local Similarity   87.5%;  Pred. No. 8.2e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TAGTTCAGGCGGGAGC 19
           |||||||| ||| |||
Db      17 TAGTTCAGCCGGTAGC 2


RESULT 12
US-08-508-088-5/c
; Sequence 5, Application US/08508088
; Patent No. 5851826
;  GENERAL INFORMATION:
;    APPLICANT:  FRAEFEL, Cornel
;    APPLICANT:  GELLER, Alfred
;    APPLICANT:  LIM, Filip
;    TITLE OF INVENTION:  HELPER VIRUS-FREE HERPESVIRUS VECTOR PACKAGING SYSTEM
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02019
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/508,088
;      FILING DATE:  26-JUL-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  45485
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;      TELEX:  200291 STRE
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
US-08-508-088-5

  Query Match             62.0%;  Score 12.4;  DB 2;  Length 24;
  Best Local Similarity   92.9%;  Pred. No. 1.2e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGTAGTTCAGGCGG 15
           |||||||||||| |
Db      24 GGTAGTTCAGGCAG 11


RESULT 13
US-09-009-925-5/c
; Sequence 5, Application US/09009925
; Patent No. 5998208
;  GENERAL INFORMATION:
;    APPLICANT:  CHILDREN'S MEDICAL CENTER CORPORATION
;    TITLE OF INVENTION:  HELPER VIRUS-FREE HERPESVIRUS VECTOR PACKAGING SYSTEM
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;      STREET:  130 WATER STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02019
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/009,925
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/12033
;      FILING DATE:  22-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/508,088
;      FILING DATE:  26-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DAVID, RESNICK S
;      REGISTRATION NUMBER:  34,235
;      REFERENCE/DOCKET NUMBER:  45485-CON
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-523-3400
;      TELEFAX:  617-523-6440
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
US-09-009-925-5

  Query Match             62.0%;  Score 12.4;  DB 2;  Length 24;
  Best Local Similarity   92.9%;  Pred. No. 1.2e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GGTAGTTCAGGCGG 15
          ||||||||||||| |
Db     24 GGTAGTTCAGGCAG 11


RESULT 14
US-09-314-847A-13/c
; Sequence 13, Application US/09314847A
; Patent No. 6365410
; GENERAL INFORMATION:
;  APPLICANT:  Schellenberger, Volker
;  APPLICANT:  Liu, Amy D.
;  APPLICANT:  Selifonova, Olga V.
;  TITLE OF INVENTION:  Directed Evolution of Microorganisms
;  FILE REFERENCE:  GC560
;  CURRENT APPLICATION NUMBER: US/09/314,847A
;  CURRENT FILING DATE:  2000-05-19
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: primer
US-09-314-847A-13

  Query Match             62.0%;  Score 12.4;  DB 4;  Length 27;
  Best Local Similarity   92.9%;  Pred. No. 1.2e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GGTAGTTCAGGCGG 15
          |||||||||| |||
Db     14 GGTAGTTCAGTCGG 1


RESULT 15
US-08-453-104-11
; Sequence 11, Application US/08453104
; Patent No. 5633446
;  GENERAL INFORMATION:
;    APPLICANT:  CORNELISSEN, Marc
;    APPLICANT:  SOETAERT, Piet
;    APPLICANT:  STAM, Maike
;    APPLICANT:  DOCKX, Jan
;    TITLE OF INVENTION:  MODIFIED BACILLUS THURINGIENSIS
;    TITLE OF INVENTION:  INSECTICIDAL - CRYSTAL PROTEIN GENES AND THEIR EXPRESSION
;    TITLE OF INVENTION:  IN PLANT CELLS
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker & Mathis
;      STREET:  George Mason Bldg., Washington & Prince Sts.
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/453,104
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/937,869
;      FILING DATE:  16-DEC-1992
;      APPLICATION NUMBER:  GB 90401055.0
;      FILING DATE:  18-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rea, Teresa S
;      REGISTRATION NUMBER:  30,427
;      REFERENCE/DOCKET NUMBER:  010830-032
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  65 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-453-104-11

  Query Match             62.0%;  Score 12.4;  DB 1;  Length 65;
  Best Local Similarity   92.9%;  Pred. No. 1.3e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 AGTTCAGGCGGGAG 18
          |||||||| |||||
Db     31 AGTTCAGGAGGGAG 44


Search completed: November 23, 2002, 06:36:05
Job time : 23.55 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-6
Perfect score: 20
Sequence: 1 gggtagttcaggcgggagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	13.2	66.0	94	10	US-09-783-590-6963	Sequence 6963, Ap
	2	12.6	63.0	36	12	US-10-001-879-17	Sequence 17, Appl
c	3	12.6	63.0	38	9	US-09-771-035A-42	Sequence 42, Appl
c	4	12.2	61.0	70	10	US-09-871-579-2	Sequence 2, Appli
c	5	12.2	61.0	70	10	US-09-871-579-14	Sequence 14, Appl
c	6	12.2	61.0	72	10	US-09-871-579-7	Sequence 7, Appli
	7	12	60.0	27	10	US-09-822-698A-15	Sequence 15, Appl
c	8	12	60.0	43	9	US-10-027-806-94	Sequence 94, Appl
	9	12	60.0	69	10	US-09-822-698A-16	Sequence 16, Appl
c	10	11.8	59.0	21	10	US-09-845-129-12	Sequence 12, Appl
c	11	11.8	59.0	33	10	US-09-828-447-29	Sequence 29, Appl
	12	11.6	58.0	21	9	US-10-032-495-40	Sequence 40, Appl
c	13	11.6	58.0	60	10	US-09-734-188-24	Sequence 24, Appl
c	14	11.6	58.0	69	10	US-09-790-417-254	Sequence 254, App
c	15	11.6	58.0	89	10	US-09-764-869-2114	Sequence 2114, Ap
c	16	11.6	58.0	89	10	US-09-764-877-3966	Sequence 3966, Ap
c	17	11.6	58.0	92	10	US-09-761-288-33	Sequence 33, Appl
c	18	11.6	58.0	92	10	US-09-761-288-37	Sequence 37, Appl
c	19	11.6	58.0	94	10	US-09-761-288-41	Sequence 41, Appl
	20	11.6	58.0	94	10	US-09-761-288-48	Sequence 48, Appl
c	21	11.6	58.0	97	10	US-09-783-590-4948	Sequence 4948, Ap
c	22	11.6	58.0	99	10	US-09-783-590-9868	Sequence 9868, Ap
c	23	11.4	57.0	18	10	US-09-969-373-3867	Sequence 3867, Ap
c	24	11.4	57.0	18	10	US-09-969-373-3869	Sequence 3869, Ap
	25	11.4	57.0	69	9	US-09-944-413-22	Sequence 22, Appl
	26	11.4	57.0	69	9	US-09-944-403-22	Sequence 22, Appl
	27	11.4	57.0	69	9	US-09-944-896-22	Sequence 22, Appl
	28	11.4	57.0	69	10	US-09-866-028-22	Sequence 22, Appl
	29	11.4	57.0	69	10	US-09-944-449-22	Sequence 22, Appl
	30	11.4	57.0	69	10	US-09-944-457-22	Sequence 22, Appl
	31	11.4	57.0	69	10	US-09-945-587-22	Sequence 22, Appl
	32	11.4	57.0	69	10	US-09-945-015-22	Sequence 22, Appl
	33	11.4	57.0	69	10	US-09-944-396-22	Sequence 22, Appl
	34	11.4	57.0	69	10	US-09-944-097-22	Sequence 22, Appl
	35	11.4	57.0	69	10	US-09-944-432-22	Sequence 22, Appl
	36	11.4	57.0	69	10	US-09-943-762-22	Sequence 22, Appl
	37	11.4	57.0	69	10	US-09-944-654-22	Sequence 22, Appl
	38	11.4	57.0	69	10	US-09-943-851A-22	Sequence 22, Appl
	39	11.4	57.0	96	10	US-09-969-373-1169	Sequence 1169, Ap
	40	11.2	56.0	18	10	US-09-969-373-2268	Sequence 2268, Ap
c	41	11.2	56.0	19	10	US-09-725-752A-11	Sequence 11, Appl
c	42	11.2	56.0	31	10	US-09-801-274-994	Sequence 994, App
c	43	11.2	56.0	89	10	US-09-764-877-2438	Sequence 2438, Ap
	44	11.2	56.0	94	10	US-09-864-761-32576	Sequence 32576, A
c	45	11.2	56.0	96	10	US-09-864-761-25922	Sequence 25922, A

ALIGNMENTS

RESULT 1
US-09-783-590-6963/c
; Sequence 6963, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6963
; LENGTH: 94
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (83)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-6963

Query Match 66.0%; Score 13.2; DB 10; Length 94;
Best Local Similarity 83.3%; Pred. No. 8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       3 GTAGTTCAGGCGGGAGCG 20
           |||| |||||| |||| |
Db      42 GTAGCTCAGGCTGGAGTG 25
```

RESULT 2
US-10-001-879-17
; Sequence 17, Application US/10001879

```
; Patent No. US20020127237A1
; GENERAL INFORMATION:
;  APPLICANT: Salceda, Susana
;  APPLICANT:  Macina, Roberto
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Cafferkey, Robert
;  APPLICANT:  Ali, Shujath
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
;  FILE REFERENCE: DEX-0281
;  CURRENT APPLICATION NUMBER: US/10/001,879
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 60/252,188
;  PRIOR FILING DATE: 2000-11-21
;  NUMBER OF SEQ ID NOS: 201
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-001-879-17

  Query Match             63.0%;  Score 12.6;  DB 12;  Length 36;
  Best Local Similarity   78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAGC 19
           ||| | | |||||| ||||
Db       4 GGGGATTACAGGCGTGAGC 22


RESULT 3
US-09-771-035A-42/c
; Sequence 42, Application US/09771035A
; Patent No. US20020165370A1
; GENERAL INFORMATION:
;  APPLICANT: KUMAGAI, Monto H.
;  APPLICANT:  ROBERTS, Peter D.
;  APPLICANT:  VAEWHONGS, Andy A.
;  TITLE OF INVENTION: CYTOPLASMIC GENE INHIBITION OR GENE
;  TITLE OF INVENTION:  EXPRESSION IN TRANSFECTED PLANTS BY TOBRAVIRAL VECTOR
;  FILE REFERENCE: 008010137CPUS04
;  CURRENT APPLICATION NUMBER: US/09/771,035A
;  CURRENT FILING DATE:  2001-01-21
;  PRIOR APPLICATION NUMBER: 09/359,301
;  PRIOR FILING DATE: 1999-07-21
;  PRIOR APPLICATION NUMBER: PCT/ US 00/20261
;  PRIOR FILING DATE: 2000-07-21
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
;   LENGTH: 38
;   TYPE: DNA
;   ORGANISM: Nicotiana benthamiana
US-09-771-035A-42

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 38;
  Best Local Similarity   78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAGC 19
           | |||||| ||||||  ||
Db      28 GCGTAGTTTAGGCGGCCGC 10


RESULT 4
US-09-871-579-2/c
; Sequence 2, Application US/09871579
; Patent No. US20020061541A1
; GENERAL INFORMATION:
;  APPLICANT: Mutter, George
;  APPLICANT:  Eng, Charis
;  TITLE OF INVENTION: Diagnosis of Endometrial Precancers
;  FILE REFERENCE: B0801/7220 (JRV)
;  CURRENT APPLICATION NUMBER: US/09/871,579
;  CURRENT FILING DATE:  2001-05-31
;  PRIOR APPLICATION NUMBER: US 60/208,542
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: US 60/289,449
;  PRIOR FILING DATE: 2001-05-08
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 70
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-871-579-2

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 70;
  Best Local Similarity   82.4%;  Pred. No. 2.4e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 TAGTTCAGGCGGGAGCG 20
           || ||| |||||| |||
Db      45 TATTTCGGGCGGGGGCG 29


RESULT 5
US-09-871-579-14/c
; Sequence 14, Application US/09871579
; Patent No. US20020061541A1
; GENERAL INFORMATION:
;  APPLICANT: Mutter, George
;  APPLICANT:  Eng, Charis
;  TITLE OF INVENTION: Diagnosis of Endometrial Precancers
;  FILE REFERENCE: B0801/7220 (JRV)
;  CURRENT APPLICATION NUMBER: US/09/871,579
;  CURRENT FILING DATE:  2001-05-31
;  PRIOR APPLICATION NUMBER: US 60/208,542
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: US 60/289,449
;  PRIOR FILING DATE: 2001-05-08
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 70
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-871-579-14

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 70;
  Best Local Similarity   82.4%;  Pred. No. 2.4e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 TAGTTCAGGCGGGAGCG 20
           || ||| |||||| |||
Db      45 TATTTCGGGCGGGGGCG 29


RESULT 6
US-09-871-579-7/c
; Sequence 7, Application US/09871579
; Patent No. US20020061541A1
; GENERAL INFORMATION:
;  APPLICANT: Mutter, George
;  APPLICANT:  Eng, Charis
;  TITLE OF INVENTION: Diagnosis of Endometrial Precancers
;  FILE REFERENCE: B0801/7220 (JRV)
;  CURRENT APPLICATION NUMBER: US/09/871,579
;  CURRENT FILING DATE:  2001-05-31
;  PRIOR APPLICATION NUMBER: US 60/208,542
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: US 60/289,449
;  PRIOR FILING DATE: 2001-05-08
;  NUMBER OF SEQ ID NOS: 20
```

```
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 72
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-871-579-7

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 72;
  Best Local Similarity   82.4%;  Pred. No. 2.4e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 TAGTTCAGGCGGGAGCG 20
          || ||| |||||| |||
Db     45 TATTTCGGGCGGGGGCG 29


RESULT 7
US-09-822-698A-15
; Sequence 15, Application US/09822698A
; Patent No. US20020146750A1
; GENERAL INFORMATION:
;  APPLICANT: Hoogenboom, Hendricus R.J.M.
;  APPLICANT:  Henderikx, Maria P.G.
;  TITLE OF INVENTION: MUCIN-1 Specific Binding Members and Methods of Use Thereof
;  FILE REFERENCE: DYX-015.1 US
;  CURRENT APPLICATION NUMBER: US/09/822,698A
;  CURRENT FILING DATE:  2001-03-30
;  PRIOR APPLICATION NUMBER: US 09/538,913
;  PRIOR FILING DATE: 2000-03-30
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: Microsoft Word
; SEQ ID NO 15
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: nucleotide sequence coding for a
;   OTHER INFORMATION: nine amino acid linker
US-09-822-698A-15

  Query Match             60.0%;  Score 12;  DB 10;  Length 27;
  Best Local Similarity   75.0%;  Pred. No. 2.9e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 GGGTAGTTCAGGCGGGAGCG 20
          |||| | ||||||||  | |
Db      3 GGGTGGATCAGGCGGCGGGG 22


RESULT 8
US-10-027-806-94/c
; Sequence 94, Application US/10027806
; Patent No. US20020160476A1
; GENERAL INFORMATION:
;  APPLICANT: Swanson, Ronald V.
;  APPLICANT:  Feldman, Robert A.
;  APPLICANT:  Schleper, Christa
;  TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
;  FILE REFERENCE: DCORP.002A
;  CURRENT APPLICATION NUMBER: US/10/027,806
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
;  NUMBER OF SEQ ID NOS: 123
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 94
;   LENGTH: 43
;   TYPE: DNA
;   ORGANISM: Cenarchaeum symbiosum
;   FEATURE:
;   NAME/KEY: TATA_signal
;   LOCATION: (11)...(16)
US-10-027-806-94

  Query Match             60.0%;  Score 12;  DB 9;  Length 43;
  Best Local Similarity   100.0%;  Pred. No. 3e+03;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 GTTCAGGCGGGA 17
          ||||||||||||
Db     28 GTTCAGGCGGGA 17


RESULT 9
US-09-822-698A-16
; Sequence 16, Application US/09822698A
; Patent No. US20020146750A1
; GENERAL INFORMATION:
;  APPLICANT: Hoogenboom, Hendricus R.J.M.
;  APPLICANT:  Henderikx, Maria P.G.
;  TITLE OF INVENTION: MUCIN-1 Specific Binding Members and Methods of Use Thereof
;  FILE REFERENCE: DYX-015.1 US
;  CURRENT APPLICATION NUMBER: US/09/822,698A
;  CURRENT FILING DATE:  2001-03-30
;  PRIOR APPLICATION NUMBER: US 09/538,913
;  PRIOR FILING DATE: 2000-03-30
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: Microsoft Word
; SEQ ID NO 16
;   LENGTH: 69
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   OTHER INFORMATION: PH1-IL-2 backward primer
US-09-822-698A-16

  Query Match             60.0%;  Score 12;  DB 10;  Length 69;
  Best Local Similarity   75.0%;  Pred. No. 3e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 GGGTAGTTCAGGCGGGAGCG 20
          |||| | ||||||||  | |
Db     24 GGGTGGATCAGGCGGCGGGG 43


RESULT 10
US-09-845-129-12/c
; Sequence 12, Application US/09845129
; Patent No. US20020146700A1
; GENERAL INFORMATION:
;  APPLICANT: DUFF, GORDON W.
;  APPLICANT:  COX, ANGELA
;  APPLICANT:  CAMP, NICOLA J.
;  APPLICANT:  DIGIOVINE, FRANCESCO S.
;  TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR DISEASES ASSOCIATED
;  TITLE OF INVENTION:  WITH AN IL-1 INFLAMMATORY HAPLOTYPE
;  FILE REFERENCE: MSA-010.02
;  CURRENT APPLICATION NUMBER: US/09/845,129
;  CURRENT FILING DATE:  2001-04-27
;  PRIOR APPLICATION NUMBER: 09/345,217
;  PRIOR FILING DATE: 1999-06-30
;  PRIOR APPLICATION NUMBER: PCT/GB98/01481
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: 9711040.7
;  PRIOR FILING DATE: 1997-05-29
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: primer
US-09-845-129-12

  Query Match             59.0%;  Score 11.8;  DB 10;  Length 21;
```

```
  Best Local Similarity   86.7%;  Pred. No. 3.7e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TAGTTCAGGCGGGAG 18
           ||||||||| || ||
Db      19 TAGTTCAGGTGGTAG 5


RESULT 11
US-09-828-447-29/c
; Sequence 29, Application US/09828447
; Patent No. US20020069432A1
; GENERAL INFORMATION:
;  APPLICANT: COSTA E SILVA, OSWALDO DA
;  APPLICANT:  BOHNERT, HANS J.
;  APPLICANT:  VAN THIELEN, NOCHA
;  APPLICANT:  CHEN, ROUYING
;  APPLICANT:  ISHITANI, MANABU
;  TITLE OF INVENTION: SIGNAL TRANSDUCTION STRESS-RELATED PROTEINS AND METHODS
;  TITLE OF INVENTION:  OF USE IN PLANTS
;  FILE REFERENCE: 16313-0037
;  CURRENT APPLICATION NUMBER: US/09/828,447
;  CURRENT FILING DATE:  2001-08-20
;  PRIOR APPLICATION NUMBER: 60/196,001
;  PRIOR FILING DATE: 2000-04-07
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-828-447-29

  Query Match             59.0%;  Score 11.8;  DB 10;  Length 33;
  Best Local Similarity   86.7%;  Pred. No. 3.7e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 AGTTCAGGCGGGAGC 19
           |||| |||| |||||
Db      19 AGTTGAGGCTGGAGC 5


RESULT 12
US-10-032-495-40
; Sequence 40, Application US/10032495
; Patent No. US20020155601A1
; GENERAL INFORMATION:
;  APPLICANT: YAN, WEN LIANG
;  TITLE OF INVENTION: METHOD FOR PRODUCING A POPULATION OF HOMOZYGOUS STEM
;  TITLE OF INVENTION:  CELLS HAVING A PRE-SELECTED IMMUNOTYPE AND/OR GENOTYPE,
;  TITLE OF INVENTION:  CELLS SUITABLE FOR TRANSPLANT DERIVED THEREFROM, AND
;  TITLE OF INVENTION:  MATERIALS AND METHODS USING SAME
;  FILE REFERENCE: 0249-0002US
;  CURRENT APPLICATION NUMBER: US/10/032,495
;  CURRENT FILING DATE:  2002-01-02
;  PRIOR APPLICATION NUMBER: 60/258,881
;  PRIOR FILING DATE: 2001-01-02
;  NUMBER OF SEQ ID NOS: 86
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-032-495-40

  Query Match             58.0%;  Score 11.6;  DB 9;  Length 21;
  Best Local Similarity   77.8%;  Pred. No. 4.6e+03;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GGTAGTTCAGGCGGGAGC 19
           || | | ||||| |||||
Db       1 GGGATTACAGGCAGGAGC 18


RESULT 13
US-09-734-188-24/c
; Sequence 24, Application US/09734188
; Patent No. US20020106645A1
; GENERAL INFORMATION:
;  APPLICANT: Richardson Ph.D., Mary Ann
;  APPLICANT:  Goldman, Assistant Counsel, Robin A.
;  APPLICANT:  New York State Office of Mental Health
;  APPLICANT:  Nathan S. Kline Institute for Psychiatric Research
;  TITLE OF INVENTION: PAH
;  FILE REFERENCE: Kline Inst.
;  CURRENT APPLICATION NUMBER: US/09/734,188
;  CURRENT FILING DATE:  2000-12-12
;  PRIOR APPLICATION NUMBER: 09/253,025
;  PRIOR FILING DATE: 2001-05-03
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-734-188-24

  Query Match             58.0%;  Score 11.6;  DB 10;  Length 60;
  Best Local Similarity   77.8%;  Pred. No. 4.7e+03;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 GTAGTTCAGGCGGGAGCG 20
           ||| | | |||||| |||
Db      46 GTAATCCGGGCGGGGGCG 29


RESULT 14
US-09-790-417-254/c
; Sequence 254, Application US/09790417
; Patent No. US20010031470A1
; GENERAL INFORMATION:
;  APPLICANT: Shultz, John W
;  APPLICANT:  Lewis, Martin K.
;  APPLICANT:  Lieppe, Donna
;  APPLICANT:  Mandrekar, Michelle
;  APPLICANT:  Kephart, Daniel
;  APPLICANT:  Rhodes, Richard B.
;  APPLICANT:  Andrews, Christine A.
;  APPLICANT:  Hartnett, James R.
;  APPLICANT:  Gu, Trent
;  APPLICANT:  Olson, Ryan J.
;  APPLICANT:  Wood, Keith W.
;  APPLICANT:  Welch, Roy
;  TITLE OF INVENTION: Nucleic Acid Detection
;  FILE REFERENCE: Pro-103 6868/75528
;  CURRENT APPLICATION NUMBER: US/09/790,417
;  CURRENT FILING DATE:  2001-02-22
;  PRIOR APPLICATION NUMBER: 09/358,972
;  PRIOR FILING DATE: 1999-07-21
;  PRIOR APPLICATION NUMBER: 09/042,287
;  PRIOR FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 290
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 254
;   LENGTH: 69
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:probe to Alu
;   OTHER INFORMATION: which forms a hairpin
US-09-790-417-254
```

```
  Query Match             58.0%;  Score 11.6;  DB 10;  Length 69;
  Best Local Similarity   77.8%;  Pred. No. 4.7e+03;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAG 18
           | || | |||||| ||||
Db      18 GTGTTGCTCAGGCTGGAG 1


RESULT 15
US-09-764-869-2114/c
; Sequence 2114, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC007
;  CURRENT APPLICATION NUMBER: US/09/764,869
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 2442
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2114
;   LENGTH: 89
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-869-2114

  Query Match             58.0%;  Score 11.6;  DB 10;  Length 89;
  Best Local Similarity   77.8%;  Pred. No. 4.7e+03;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 GTAGTTCAGGCGGGAGCG 20
           ||  ||||||| |||| |
Db      36 GTCATTCAGGCTGGAGTG 19


Search completed: November 23, 2002, 06:42:05
Job time : 18.25 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-6
Perfect score: 20
Sequence: 1 gggtagttcaggcgggagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	97	9	AA406204	AA406204 zu65d04.s
	2	14.2	71.0	97	17	CNS02VXN	AL216356 Tetraodon
c	3	13.8	69.0	99	14	T72387	T72387 yc71d12.r1
c	4	13.6	68.0	32	17	AZ377853	AZ377853 1M0132E21
c	5	13.6	68.0	61	17	AZ920486	AZ920486 1006020A0
c	6	13.6	68.0	92	17	BH412029	BH412029 1007024H0
c	7	13.2	66.0	38	17	AZ440246	AZ440246 1M0231A18
	8	13.2	66.0	51	9	AI033458	AI033458 ow94h01.s
c	9	13.2	66.0	51	13	BM431597	BM431597 1Duo24E12
c	10	13.2	66.0	60	17	CNS02V1M	AL215203 Tetraodon
	11	13.2	66.0	65	17	CNS02VVD	AL216274 Tetraodon
	12	12.8	64.0	23	13	BM398474	BM398474 5009-0-45
	13	12.8	64.0	25	14	L32042	L32042 HUMXP2E8A H
c	14	12.8	64.0	42	17	AZ650453	AZ650453 1M0520F10
	15	12.8	64.0	77	17	TA126F03Q	AL463415 T. brucei
c	16	12.8	64.0	95	14	BQ821804	BQ821804 1030095B1
	17	12.6	63.0	50	9	AU104217	AU104217 AU104217
	18	12.6	63.0	69	14	BQ761873	BQ761873 EBpi01_SQ
c	19	12.6	63.0	77	10	AW022596	AW022596 df41b06.y
c	20	12.6	63.0	77	17	AZ477146	AZ477146 1M0296B01
	21	12.6	63.0	80	17	BH215982	BH215982 1006039C1
	22	12.6	63.0	85	14	H96969	H96969 yw11d11.s1
c	23	12.6	63.0	88	10	AW302675	AW302675 xr55d12.x
c	24	12.6	63.0	100	9	AI418498	AI418498 tf69e03.x
	25	12.6	63.0	100	12	BE909180	BE909180 601501419
	26	12.4	62.0	58	17	BH792160	BH792160 SALK_0628
	27	12.4	62.0	68	14	T52860	T52860 ya81d02.r1
c	28	12.2	61.0	41	14	T85260	T85260 yd45g03.s1
c	29	12.2	61.0	48	17	AZ326039	AZ326039 1M0048N18
c	30	12.2	61.0	50	9	AU104765	AU104765 AU104765
	31	12.2	61.0	60	10	AW691683	AW691683 NF048A05S
c	32	12.2	61.0	64	12	BF228772	BF228772 SWOvL3CAN
c	33	12.2	61.0	71	9	AI904673	AI904673 QV-BT065-
	34	12.2	61.0	76	17	BH188061	BH188061 035_C_06-
	35	12.2	61.0	76	17	CNS07S62	AL625004 T3 end of
	36	12.2	61.0	85	17	BH408656	BH408656 1007006E1
c	37	12.2	61.0	87	9	AA229989	AA229989 nc37d09.s
	38	12.2	61.0	92	9	AA027368	AA027368 mi04d06.r
c	39	12.2	61.0	95	17	BH218367	BH218367 1006078E0
c	40	12.2	61.0	97	14	R51054	R51054 yg71g08.s1
c	41	12.2	61.0	97	14	R72811	R72811 yl09g09.r1
c	42	12.2	61.0	98	17	CNS0429N	AL271220 Tetraodon
c	43	12	60.0	25	9	AI913416	AI913416 tz77a09.x
c	44	12	60.0	28	14	N81476	N81476 TgESTzy47h0
c	45	12	60.0	46	9	AI809473	AI809473 wh76c10.x

ALIGNMENTS

RESULT 1
AA406204/c
LOCUS AA406204 97 bp mRNA linear EST 17-MAY-1997
DEFINITION zu65d04.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:742855 3' similar to SW:NUKM_BOVIN P42026 NADH-UBIQUINONE OXIDOREDUCTASE 20 KD SUBUNIT PRECURSOR ;, mRNA sequence.
ACCESSION AA406204
VERSION AA406204.1 GI:2064185
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 97)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand

Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .97
/organism="Homo sapiens"
/db_xref="GDB:5929850"
/db_xref="taxon:9606"
/clone="IMAGE:742855"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories , Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 15 a 33 c 27 g 22 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 9; Length 97;
Best Local Similarity 84.2%; Pred. No. 2.7e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 GGTAGTTCAGGCGGGAGCG 20
           || ||  ||| |||||||||
Db      42 GGAAGATCAAGCGGGAGCG 24
```

RESULT 2
CNS02VXN
LOCUS CNS02VXN 97 bp DNA linear GSS 15-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 175B09 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL216356
VERSION AL216356.1 GI:7875175
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 97)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 97)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Charaterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 97)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
FEATURES Location/Qualifiers
source 1. .97
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="175B09"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG175CA05LP1~end : T7"
BASE COUNT 22 a 31 c 29 g 13 t 2 others
ORIGIN

Query Match 71.0%; Score 14.2; DB 17; Length 97;
Best Local Similarity 84.2%; Pred. No. 2.7e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 GGTAGTTCAGGCGGGAGCG 20
           ||| |||||||| |||| |
Db      55 GGTCGTTCAGGCAGGAGGG 73
```

RESULT 3
T72387/c
LOCUS T72387 99 bp mRNA linear EST 01-MAR-1995
DEFINITION yc71d12.r1 Stratagene liver (#937224) Homo sapiens cDNA clone IMAGE:86135 5' similar to gb:X06985_rna1 HEME OXYGENASE 1 (HUMAN);, mRNA sequence.
ACCESSION T72387
VERSION T72387.1 GI:686908
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 99)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 43
High qality sequence starts: 1 High qality sequence stops: 1
Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality
Seq primer: M13RP1
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .99
/organism="Homo sapiens"
/db_xref="GDB:503192"
/db_xref="taxon:9606"
/clone="IMAGE:86135"
/clone_lib="Stratagene liver (#937224)"
/sex="male"
/dev_stage="49 years old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: liver; Vector: pBluescript SK; Site_1: EcoRI ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Hepatectomy from normal male caucasian. Average insert size: 1.1 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT 23 a 29 c 23 g 15 t 9 others
ORIGIN

Query Match 69.0%; Score 13.8; DB 14; Length 99;
Best Local Similarity 78.9%; Pred. No. 4e+04;

```
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GGTAGTTCAGGCGGGAGCG 20
          || |||||| |||| |||
Db     49 GGGAGTTCATGCGGNNGCG 31


RESULT 4
AZ377853/c
LOCUS       AZ377853                 32 bp    DNA     linear   GSS 02-OCT-2000
DEFINITION  1M0132E21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0132E21 F, DNA sequence.
ACCESSION   AZ377853
VERSION     AZ377853.1  GI:10491553
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0132  row: E  column: 21
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 32.
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0132E21"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        7 a     14 c      4 g      7 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 32;
  Best Local Similarity   80.0%;  Pred. No. 4.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GGGTAGTTCAGGCGGGAGCG 20
          || ||   ||||||||||||
Db     20 GGATAAGCCAGGCGGGAGCG 1


RESULT 5
AZ920486/c
LOCUS       AZ920486                 61 bp    DNA     linear   GSS 17-DEC-2001
DEFINITION  1006020A02.x1 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   AZ920486
VERSION     AZ920486.1  GI:13391189
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Sequence was trimmed at very probable ligation site. Post-ligation
            sequence submitted separately.
            Plate: 1006020  row: 38
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT        6 a     27 c     19 g      9 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 61;
  Best Local Similarity   80.0%;  Pred. No. 4.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GGGTAGTTCAGGCGGGAGCG 20
          ||||||||   ||||| |||
Db     61 GGGTAGTTTGCGCGGGCGCG 42


RESULT 6
BH412029/c
LOCUS       BH412029                 92 bp    DNA     linear   GSS 12-DEC-2001
DEFINITION  1007024H03.y1 1007 - RescueMu Grid H Zea mays genomic, DNA
            sequence.
```

```
ACCESSION   BH412029
VERSION     BH412029.1  GI:17584839
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site of ends cut by single endonuclease.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1007024  column: 14
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1007 - RescueMu Grid H"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units.  For more information on RescueMu, go to the  web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.'  Grid H was grown at Berkeley in 2001.  DNA
                     was extracted from leaf punches,  double digested using
                     BamHI and BglII, and ligated to form circular plasmids.
                     DH10B cells were transformed and then screened on LB
                     plates with ampicillin."
BASE COUNT       13 a     41 c     26 g     12 t
ORIGIN

  Query Match           68.0%;  Score 13.6;  DB 17;  Length 92;
  Best Local Similarity   80.0%;  Pred. No. 4.8e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GGGTAGTTCAGGCGGGAGCG 20
           |||   |||| |||||||||
Db      49 GGGGTCTTCAAGCGGGAGCG 30


RESULT 7
AZ440246/c
LOCUS       AZ440246                38 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0231A18F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0231A18 F, DNA sequence.
ACCESSION   AZ440246
VERSION     AZ440246.1  GI:10564259
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0231  row: A  column: 18
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 38.
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0231A18"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        9 a     14 c      5 g     10 t
ORIGIN

  Query Match           66.0%;  Score 13.2;  DB 17;  Length 38;
  Best Local Similarity   83.3%;  Pred. No. 6.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGTAGTTCAGGCGGGAGC 19
           |||| || |||| |||||
Db      30 GGTAATTTAGGCAGGAGC 13


RESULT 8
AI033458
LOCUS       AI033458                51 bp    mRNA    linear   EST 28-AUG-1998
DEFINITION  ow94h01.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:1654513 3' similar to gb:M81181
            SODIUM/POTASSIUM-TRANSPORTING ATPASE BETA-2 (HUMAN);, mRNA
            sequence.
ACCESSION   AI033458
VERSION     AI033458.1  GI:3254411
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
```

```
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 400   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1654513"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.   1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer {5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'},
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT        6 a      7 c     17 g     21 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 6.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 GTAGTTCAGGCGGGAGCG 20
          |||| |||||| |||| |
Db     33 GTAGGTCAGGCTGGAGTG 50


RESULT 9
BM431597/c
LOCUS       BM431597                 51 bp    mRNA    linear   EST 31-JAN-2002
DEFINITION  1Duo24E12.ab1 Bos taurus Duodenum #1 library Bos taurus cDNA, mRNA
            sequence.
ACCESSION   BM431597
VERSION     BM431597.1  GI:18453319
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Hansen,C., Fu,A., Meng,Y., Li,C., Okine,E., Sensen,C.W., Gordon
            ,P.M.K. and Moore,S.S.
  TITLE     Gene Expression Profiling of the Bovine Gastrointestinal Tract
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Dr. Stephen Moore
            . Beef Genomics Laboratory
            Dept of AFNS, University of Alberta
            410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
            Tel: 780 492 0169
            Fax: 780 492 4265
            Email: smoore@afns.ualberta.ca
            Insert Length: 51   Std Error: 0.00
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="Bos taurus Duodenum #1 library"
                     /tissue_type="Smooth muscle"
                     /cell_type="Simple columnar epithelial"
                     /dev_stage="Young adult"
                     /lab_host="XL1-BlueMRF'strain"
                     /note="Organ: Intestine/duodenum; Vector: Uni-2ZAPXR;
                     Site_1: EcoRI; Site_2: Xho I"
BASE COUNT        9 a     21 c     13 g      8 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 13;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 6.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGTAGTTCAGGCGGGAGC 19
          ||  ||||||||| ||||
Db     27 GGGGGTTCAGGCGCGAGC 10


RESULT 10
CNS02V1M/c
LOCUS       CNS02V1M                 60 bp    DNA     linear   GSS 15-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            173I24 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL215203
VERSION     AL215203.1  GI:7874022
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 60)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 60)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="173I24"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG173BE12SP1-end :
                     PUC-Ori"
BASE COUNT        7 a     25 c     13 g     14 t      1 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 60;
  Best Local Similarity   83.3%;  Pred. No. 6.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGTAGTTCAGGCGGGAGC 19
          ||||| | || |||||||
```

```
Db      40 GGTAGCTGAGACGGGAGC 23


RESULT 11
CNS02VVD
LOCUS       CNS02VVD                  65 bp    DNA     linear   GSS 15-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            175G08 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL216274
VERSION     AL216274.1  GI:7875093
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 65)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 65)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="175G08"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG175BD04LP1-end : T7"
BASE COUNT       14 a     16 c     30 g      4 t      1 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 65;
  Best Local Similarity   83.3%;  Pred. No. 6.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGTAGTTCAGGCGGGAGC 19
           ||||| | || |||||||
Db       6 GGTAGCTGAGACGGGAGC 23


RESULT 12
BM398474
LOCUS       BM398474                  23 bp    mRNA    linear   EST 17-JAN-2002
DEFINITION  5009-0-45-G04.t.1 Chilcoat/Turkewitz cDNA (large fraction)
            Tetrahymena thermophila cDNA, mRNA sequence.
ACCESSION   BM398474
VERSION     BM398474.1  GI:18198527
KEYWORDS    EST.
SOURCE      Tetrahymena thermophila.
  ORGANISM  Tetrahymena thermophila
            Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
            Hymenostomatida; Tetrahymenina; Tetrahymena.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Turkewitz,A.P., Karrer,K.M., Jahn,C., Orias,E., Kirk,K.E., Frankel
            ,J. and Klobutcher,L.
  TITLE     EST from Tetrahymena thermophila, strain CU428.1, growing cells
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Turkewitz AP
            Molecular Genetics and Cell Biology
            University of Chicago
            920 E. 58th Street, Chicago, IL 60637, USA
            Tel: 773 702 4374
            Fax: 773 702 3172
            Email: apturkew@midway.uchicago.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="Tetrahymena thermophila"
                     /strain="CU428.1"
                     /db_xref="taxon:5911"
                     /clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
                     /note="Vector: BlueScript2 SK+; Details on library
                     preparation can be found in Chilcoat and Turkewitz (2001)
                     Proc. Natl. Acad. Sci USA, 98: 8709-8713."
BASE COUNT        3 a      7 c      9 g      4 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 13;  Length 23;
  Best Local Similarity   87.5%;  Pred. No. 9e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 AGTTCAGGCGGGAGCG 20
           || ||| |||||||||
Db       4 AGCTCACGCGGGAGCG 19


RESULT 13
L32042
LOCUS       L32042                    25 bp    mRNA    linear   EST 02-AUG-1995
DEFINITION  HUMXP2E8A Human placenta Homo sapiens cDNA clone XP2E8A, mRNA
            sequence.
ACCESSION   L32042
VERSION     L32042.1  GI:557153
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Lee,C.-C., Yazdani,A., Wehnert,M., Bailey,J., Couch,L., Xiong,M.,
            Coolbaugh,M.I., Chinault,C.A., Baldini,A., Lindsay,E.A., Zhao,Z.-Y.
            and Caskey,C.T.H.
  TITLE     Isolation of chromosome-specific genes by reciprocal probing of
            arrayed cDNAs and cosmid libraries
  JOURNAL   Hum. Mol. Genet. 4, 1373-1380 (1995)
  MEDLINE   96090257
COMMENT     Contact: Caskey,C.T.H.
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="Xp22.1"
                     /clone="XP2E8A"
                     /clone_lib="Human placenta"
                     /note="Arrayed cDNAs and cosmid libraries from human
                     placental tissue"
BASE COUNT        6 a      4 c     10 g      5 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 14;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 9.1e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GGTAGTTCAGGCGGGA 17
           |||| ||||||| |||
Db       1 GGTATTTCAGGCAGGA 16
```

```
RESULT 14
AZ650453/c
LOCUS       AZ650453                 42 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0520F10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0520F10 R, DNA sequence.
ACCESSION   AZ650453
VERSION     AZ650453.1  GI:11784957
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0520  row: F  column: 10
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 42.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0520F10"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       10 a     15 c      5 g     12 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 42;
  Best Local Similarity   87.5%;  Pred. No. 9.7e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GGTAGTTCAGGCGGGA 17
            ||||||||||  ||||
Db      36 GGTAGTTCAGAGGGGA 21
```

```
RESULT 15
TA126F03Q
LOCUS       TA126F03Q                77 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 126f03, reverse sequence,
            genomic survey sequence.
ACCESSION   AL463415
VERSION     AL463415.1  GI:11833925
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="126f03"
BASE COUNT       30 a     12 c     18 g     17 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 77;
  Best Local Similarity   87.5%;  Pred. No. 1e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 AGTTCAGGCGGGAGCG 20
            |||| |||||||||| |
Db      51 AGTTAAGGCGGGAGGG 66


Search completed: November 26, 2002, 04:07:14
Job time : 766.8 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-7
Perfect score: 20
Sequence: 1 aatggcgccctgtgtcccga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		14.2	71.0	25	6	AR184007	AR184007 Sequence
2		14.2	71.0	25	6	I68645	I68645 Sequence 10
3		13.8	69.0	44	6	AX027575	AX027575 Sequence
4	c	13.8	69.0	45	6	AX027576	AX027576 Sequence
5	c	13.6	68.0	30	6	AX278367	AX278367 Sequence
6		13.4	67.0	78	3	AF090124	AF090124 Helix asp
7	c	13.2	66.0	20	6	AR163949	AR163949 Sequence
8	c	13.2	66.0	20	6	AX326934	AX326934 Sequence
9		13.2	66.0	34	6	A60217	A60217 Sequence 13
10		13.2	66.0	34	6	AR122293	AR122293 Sequence
11	c	13.2	66.0	40	6	AR135209	AR135209 Sequence
12	c	13.2	66.0	40	6	AR146705	AR146705 Sequence
13	c	13.2	66.0	40	6	AR152276	AR152276 Sequence
14	c	13.2	66.0	40	6	AR157814	AR157814 Sequence
15		13.2	66.0	45	9	HSA403877	AJ403877 Homo sapi
16		13.2	66.0	77	9	HSRGS08	AF030141 Homo sapi
17	c	13.2	66.0	79	6	A50092	A50092 Sequence 44
18	c	13	65.0	31	6	AX248740	AX248740 Sequence
19	c	12.8	64.0	21	6	AR120005	AR120005 Sequence
20		12.8	64.0	21	6	AR120068	AR120068 Sequence
21	c	12.8	64.0	21	6	I13801	I13801 Sequence 9
22		12.8	64.0	21	6	I13864	I13864 Sequence 72
23	c	12.8	64.0	21	6	I33435	I33435 Sequence 9
24	c	12.8	64.0	53	6	A84386	A84386 Sequence 1
25	c	12.8	64.0	53	6	AX427285	AX427285 Sequence
26	c	12.8	64.0	78	3	URDRRE05	M81027 Urodacus ho
27	c	12.6	63.0	65	6	AR064497	AR064497 Sequence
28		12.6	63.0	72	9	S80783	S80783 mannose 6-p
29		12.6	63.0	72	9	S80785	S80785 mannose 6-p
30	c	12.4	62.0	51	6	AX160189	AX160189 Sequence
31	c	12.4	62.0	51	6	AX165134	AX165134 Sequence
32	c	12.2	61.0	20	6	AR136418	AR136418 Sequence
33	c	12.2	61.0	25	6	AX196730	AX196730 Sequence
34		12.2	61.0	28	6	AX040164	AX040164 Sequence
35	c	12.2	61.0	30	10	MMHNF1I3B	X66552 Mus musculu
36	c	12.2	61.0	32	9	S81153	S81153 T cell anti
37		12.2	61.0	39	6	AX097394	AX097394 Sequence
38		12.2	61.0	40	6	I07950	I07950 Sequence 3
39	c	12.2	61.0	41	6	AX135990	AX135990 Sequence
40	c	12.2	61.0	41	6	AX136047	AX136047 Sequence
41	c	12.2	61.0	41	6	BD006882	BD006882 Oligonucl
42		12.2	61.0	45	6	AX027573	AX027573 Sequence
43	c	12.2	61.0	45	6	AX027574	AX027574 Sequence
44	c	12.2	61.0	45	10	MMF12RNA	X60843 M.musculus
45	c	12.2	61.0	51	6	AX157611	AX157611 Sequence

ALIGNMENTS

RESULT 1
AR184007
LOCUS AR184007 25 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 10 from patent US 6342483.
ACCESSION AR184007
VERSION AR184007.1 GI:20227976
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 25)
AUTHORS Holt,J.T., Jensen,R.A., Page,D.L., Obermiller,P.S., Robinson-Benion,C.L. and Thompson,M.E.
TITLE Method for detection and treatment of breast cancer
JOURNAL Patent: US 6342483-A 10 29-JAN-2002;

```
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="unknown"
BASE COUNT        4 a      8 c      8 g      5 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 25;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ATGGCGCCCTGTGTCCCGA 20
          |||||| |||||  |||||
Db      3 ATGGCGTCCTGTAACCCGA 21


RESULT 2
I68645
LOCUS       I68645                    25 bp    DNA     linear   PAT 04-FEB-1998
DEFINITION  Sequence 10 from patent US 5677125.
ACCESSION   I68645
VERSION     I68645.1  GI:2830767
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Holt,J.T., Jensen,R.A., Page,D.L., Obermiller,P.S. and
            Robinson-Benion,C.L.
  TITLE     Method of detection and diagnosis of pre-invasive cancer
  JOURNAL   Patent: US 5677125-A 10 14-OCT-1997;
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="unknown"
BASE COUNT        4 a      8 c      8 g      5 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 25;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ATGGCGCCCTGTGTCCCGA 20
          |||||| |||||  |||||
Db      3 ATGGCGTCCTGTAACCCGA 21


RESULT 3
AX027575
LOCUS       AX027575                  44 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 35 from Patent WO0039308.
ACCESSION   AX027575
VERSION     AX027575.1  GI:10188483
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Sakai,Y., Hashimoto,Y. and Takemoto,Y.
  TITLE     Ikk3 kinase
  JOURNAL   Patent: WO 0039308-A 35 06-JUL-2000;
            SAKAI YUTAKA (JP) ; GLAXO WELLCOME KABUSHIKI KAISH (JP) ; HASHIMOTO
            YASUHIRO (JP) ; TAKEMOTO YOSHIHIRO (JP)
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer"
BASE COUNT        6 a      14 c     18 g      6 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 44;
  Best Local Similarity   88.2%;  Pred. No. 3.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GGCGCCCTGTGTCCCGA 20
          ||||||||| |||| ||
Db     24 GGCGCCCTGGGTCCGGA 40


RESULT 4
AX027576/c
LOCUS       AX027576                  45 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 36 from Patent WO0039308.
ACCESSION   AX027576
VERSION     AX027576.1  GI:10188484
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Sakai,Y., Hashimoto,Y. and Takemoto,Y.
  TITLE     Ikk3 kinase
  JOURNAL   Patent: WO 0039308-A 36 06-JUL-2000;
            SAKAI YUTAKA (JP) ; GLAXO WELLCOME KABUSHIKI KAISH (JP) ; HASHIMOTO
            YASUHIRO (JP) ; TAKEMOTO YOSHIHIRO (JP)
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer"
BASE COUNT        6 a      19 c     14 g      6 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 45;
  Best Local Similarity   88.2%;  Pred. No. 3.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GGCGCCCTGTGTCCCGA 20
          ||||||||| |||| ||
Db     21 GGCGCCCTGGGTCCGGA 5


RESULT 5
AX278367/c
LOCUS       AX278367                  30 bp    DNA     linear   PAT 01-NOV-2001
DEFINITION  Sequence 30 from Patent WO0177169.
ACCESSION   AX278367
VERSION     AX278367.1  GI:16605376
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Tang,T.Y., Labat,I., Tillinghast,J.S., Sinku,A., Liu,C.,
            Drmanac,R.T., Stache-Crain,B., Dickson,M., Mize,N.K. and
            Nishikawa,M.
  TITLE     Methods and materials relating to novel stem cell growth
            factor-like polypeptides and polynucleotides
  JOURNAL   Patent: WO 0177169-A 30 18-OCT-2001;
            HYSEQ, INC. (US) ; KIRIN BEER KABUSHIKI KAISHA (JP)
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="PCR primer"
BASE COUNT        5 a      6 c      11 g     8 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 4.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 AATGGCGCCCTGTGTCCCGA 20
          |||||||| ||   ||||||
Db     21 AATGGCGCGCTTCTTCCCGA 2
```